Exploring general outcome measures for decision making using a wearable eye tracker

Tiemen Wagenvoort, 5691958, t.j.wagenvoort@uu.nl

Applied Cognitive Psychology

Thesis (combined with internship), 27.5 ETCS

Utrecht University

First assessor: Jeroen Benjamins, j.s.benjamins@uu.nl

Second assessor: Ignace Hooge, i.hooge@uu.nl

03/07/2020. Public show: yes.

Abstract

Studying gaze behavior using eye tracking can give insights in decision making processes regarding food choices. However, commonly used analyses that involve area of interests (AOI's) are difficult to implement in measurements made in complex environments like a supermarket. In these cases it could be beneficial to use general outcome measures that do not use AOI's. In this study, we investigated with a wearable eye tracker if general outcome measures can account for differences in gaze behavior for customers that already decided what to buy (resembling a search task) or did not yet decided what to buy (resembling a decision task). Additionally, we took into account their familiarity with the supermarket (resembling learning effects). This study substantiates the Natural Decision Segmentation Model by providing evidence for differences in gaze behavior during the 'orientation' compared to the 'evaluation and verification' phase using general outcome measures. More specifically, undecided unfamiliar customers had higher standard deviations of saccadic amplitude and velocity, and longer mean saccadic durations in the orientation phase compared to other customers. Although the results should be considered with nuance, they seem to support the feasibility of wearable eye tracking research in complex and applied settings by means of general outcome measures analyses.

Keywords: *wearable eye tracking, food choices, decision phases, search phases, general outcome measures*

Introduction

During a typical visit to the grocery store we are usually faced with a large amount of decisions. These decisions involve, among other things, which products to choose and how to navigate our bodies through the environment of the supermarket. Over the years, several theories have emerged on how we make decisions. It is often assumed that all information relevant to a choice is attended to and processed (March, 1978; Simon, 1955). However, limits in cognitive capacity challenge this theory of full rationality (Gigerenzer & Gaissmaier, 2011; Simon, 1955), suggesting that decision makers use different strategies to reach their conclusions effectively. These strategies are likely to vary depending on several factors, such as when someone is familiar or not with the environment in which the decision has to be made (Park & Lessig, 1981; Hoyer, 1984), or when the decision maker did, or did not, already decided what to choose when entering this environment (Ajzen, 1991).

An increasingly popular method to investigate the cognitive processes underlying decision making is eye tracking (Orquin & Loose, 2013). With eye tracking, gaze behavior can be classified as events such as fixations and saccades. Although these concepts are widely used, there appears to be no universal, formal definitions among vision scientists around the world (Hessels, Niehorster, Nyström, Andersson, & Hooge, 2018). Here, we define fixations as relative slow phases in the signal of a wearable eye tracker as classified by the algorithm of Hessels, van Doorn, Benjamins, Holleman, and Hooge (2020) with default settings, often serving to maintain a certain area of, or object in, the visual field at a relative constant location on the retina. Notice that this definition of a fixation allows for some degree of eye movement in relation to the head. Saccades are defined as relative fast phases according to the same algorithm, often serving to bring a new area of, or object in, the visual field to the fovea. Attention is usually located at the point of fixation (i.e. overt attention; Orquin & Loose, 2013), thereby opening up an opportunity to apply eye tracking in decision making research. For example, when certain features of the environment are more salient than others (Itti & Koch, 2001), it is more likely that these features are fixated and cognitively processed (Lohse, 1997; Milosavljevic, Navalpakkam, Koch, & Rangel, 2012). This bottom-up attention contrasts with top-down attention, when features are more likely to be fixated because of internal processes such as the memory or preferences of the observer (Glaholt, Wu, & Reingold, 2010; Van Herpen & Van Trijp, 2011). Usually, bottom-up and top-down attention will interact and subsequently influence gaze behavior equally (Van der Stigchel et al., 2013; Gidlöf, Anikin, Lingonblad, & Wallin, 2017), but there are exceptions. For example, top-down control of eye movements increases with increased experience. Decision makers subsequently attend to high utility features more often (Orquin, Bagger, & Loose, 2013). Similarly, this study found that the influence of presentation format on attention capture reduced over time, suggesting that bottom-up control of eye movements decreased.

This tight interplay between vision and attention (i.e. eye-mind hypothesis) causes eye tracking to be a fruitful method when researching food decision making. For example, Danner et al. (2016) found that more fixations at a food product, a higher amount of dwells (i.e. visits in an area of interest (AOI), from entry to exit (Holmqvist et al., 2011)), and longer total dwell time (cumulative dwell time of an AOI over the entire trial) were related to an increased likelihood of choice. Furthermore, Peschel and Orquin (2013) created a model to explain why larger surface sizes attract more fixations and are fixated faster than small objects, and subsequently increase the likelihood of choice. They conclude that this surface size effect is bottom-up, resulting from object size, the amount of objects in the environment and object distance to the center of the visual scene. Gidlöf et al. (2017) show support for the relevance of these external, bottom-up factors, but also stress the role of internal, top-down processes such as consumer preference, by using a wearable eye tracker to investigate food choices in a supermarket. Additionally, they highlight the importance of visual attention for actual purchases, even after controlling for all external and internal factors. This suggests that the likelihood of choice increases just because a product is being gazed at more often or longer. There is considerable debate about this downstream effect of vision on choice (Orquin & Loose, 2013). The attentional drift diffusion model, for example, indicates that more visual attention causes the attended item to be more likely to be chosen (Krajbich & Rangel, 2011). Van der Laan, Hooge, De Ridder, Viergever, and Smeets (2015) disentangled the decision goal (task instruction) and preference formation to research to role of total fixation duration on a product. Even though total fixation duration was mostly influenced by the decision goal (in line with the seminal work of Yarbus (1967)), preference formation also contributed. Orquin and Loose (2013) attribute the relation between increased visual attention and increased choice likelihood to the creation of consideration sets. They state that by looking at a product its information can subsequently be considered, whereas not looking at a product makes this challenging.

Two observations should be made about the aforementioned studies and others in the field of eye tracking and decision making. First, technological advances have made it possible to study gaze behavior in more complex environments like a supermarket besides the classic lab studies. Although many principles regarding vision and choice will be the same in both settings, challenges can arise in the methods that are used (Hessels et al., 2020; Orquin & Holmqvist, 2018). Second, using AOI's is common practice when analyzing eye tracking data. For example, dwells, total fixation duration and other AOI measures are often used, but the debate on the downstream effect of vision on choice gives an example on difficulties to interpret these measures. Together, these two observations create several challenges.

First of all, it can be problematic to create meaningful AOI's. When studying gaze in a supermarket, for example, it is straightforward to classify several product groups, signs, other customers or features of the environment as AOI's. Problems arise, however, when the chosen product, for instance an apple, is classified as an AOI as well. This apple could be classified as both the chosen product and the product group 'fruit', thus creating overlapping AOI's, which is problematic for further analysis (Orquin, Ashy, & Clarke, 2016). Creating AOI's can also be a problem because often there is no consensus how to define an AOI or a certain stimulus, even when the researchers operate within the same field (Hessels, Kemner, van den Boomen, & Hooge (2016). Furthermore, it can be problematic to reliably assign fixations to AOI's. Certain eye trackers lack precision, or their precision varies over the trial (Holmqvist et al., 2011; Holmqvist, Nyström, & Mulvey, 2012). This is especially an issue when investigating complex environments such as a grocery store, in which small products (i.e. small AOI's) are often placed next to other small products.

When using wearable eye trackers, software like Gazecode (Benjamins, Hessels, & Hooge, 2018) and GlassesViewer (Niehorster, Hessels, & Benjamins, 2019) allow for manual classification and data quality checks of wearable eye tracker data, but it remains challenging to classify fixations to objects reliably in complex environments (Orquin et al., 2016). Additionally, even if it is possible to create meaningful AOI's and classify fixations properly, it can be problematic to interpret the results of the analyses for three reasons. First, attention is not necessarily located at the point of fixation, but it can be located in the periphery as well (i.e. covert attention; Posner, 1980). Usually this only occurs just prior to a saccade (Shepherd, Findlay & Hockey, 1986), but it can also be done voluntarily. When only using the eye tracking data, it is ambiguous if participants actually cognitively processed an item when fixating it, since the eye tracking data does not show when they attended overtly or covertly. Similarly, it is uncertain if participants did not process an item when they did not fixated it (Van Loo, Nayga, Campbell, Seo, & Verbeke, 2018). Second, gaze behavior is not necessarily related to cognitive processes, but it is indicative of physical processes, such as navigating the environment (Hayhoe & Ballard, 2005, Hessels et al., 2020) or reaching for an item, as well (Johansson, Westling, Bäckström & Flanagan, 2001; Land & Hayhoe, 2001). For example, when classifying dwell time for a chosen product, it is impossible to know whether participants are fixating the product because they are cognitively processing it, or whether they are fixating it because they prepare or execute a physical movement to walk towards the product or to grab it (i.e. hand-eye coordination). Third, even if it would be assumed that a stimulus within an AOI is cognitively processed when fixated, it is challenging to identify which cognitive process is at play (Orquin & Holmqvist, 2018). This is for instance shown by the discussion on the downstream effect of vision on choice (Orquin & Loose, 2013), as mentioned earlier, or by the interaction between bottom-up and top-down attention (Gidlöf et al., 2017). For an elaborate explanation on total dwell time regarding this issue, please refer to Orquin & Holmqvist (2018).

One way to partially avoid these type of problems is to not use eye movement measures that are directly related to objects. Instead of analyzing *what* participants looked at, it could be analyzed *how* they looked during the experiment, irrespective of the content of the visual stimulus. This includes measures such as mean fixation duration (regardless of what is fixated) and saccadic amplitude, duration and velocity during a certain viewing period. Notice that these general outcome measures still include fixations and saccades, but that they are used in a less conventional way by not assigning them to a stimulus or object within AOI's. Although it might feel counterintuitive to decouple eye tracking data from the visual stimuli when using global outcome measures, there are several studies providing support for the usefulness of these measures. For example, Over, Hooge, Vlaskamp, and Erkelens (2007) showed how mean saccadic amplitude decreased and mean fixation duration increased gradually as participants searched for features in a cluttered scene longer. These global outcome measures did, except for the target, not make use of AOI's but still provided evidence for a course-to-fine eye movement strategy in visual search. Similarly, there appear to be different phases in attention during decision making which can be identified with eye tracking (Orquin & Loose, 2013; Russo & Leclerc, 1994). There is no consensus on the exact amount of phases and their characteristics yet, but generally a distinction is made between 'orientation' or 'scanning', 'comparison' or ‘evaluation’, and ‘verification’. Glöckner and Herbold (2011) found that the first 10 to 20 fixations in the first phase contain a larger share of short fixation durations (< 150 ms) than the other phases, showing that general outcome measures can be indicative of decision phases as well. Gidlöf, Wallin, Dewhurst, and Holmqvist (2013) noted that these attention phases might represent search strategies rather than decision making processes, and therefore compared participants who either were instructed to find a specific product in a supermarket (search task) or to buy a product of their own choice (decision task). Meanwhile, they recorded their gaze behavior using a wearable eye tracker. They conclude that their Natural Decision Segmentation Model (NDSM) identifies differences between the search and decision task better than the classical approach from Russo and Leclerc (1994). A key feature seems to be the higher amount of re-fixations or re-dwells in the second phase (‘comparison’) during a decision task compared to a search task. Notice that defining re-fixations or re-dwells requires the use of AOI’s, so general outcome measures were not used.

In this study, we tested with a wearable eye tracker (1) if general outcome measures can account for differences in gaze behavior for customers that already decided what to buy (resembling a search task) or did not yet decided what to buy (resembling a decision task) when entering a supermarket. Additionally, we investigated if this interacts with the amount of familiarity the customer has with this particular supermarket (resembling learning effects). Lastly, (2) we used general outcome measures to identify phases in decision making in an exploratory fashion. Based on Ajzen (1991) and Orquin & Loose (2013), we hypothesized that customers who already decided what to buy have a smaller average saccadic amplitude, shorter average saccadic duration, slower average saccadic velocity and lower standard deviations of saccadic amplitude, duration and velocity than customers who decided what to buy while being in the supermarket. We expected that these effects are amplified if someone is familiar with the supermarket (i.e. visits it often), compared to when someone is less so (Park & Lessig, 1981; Hoyer, 1984; Orquin & Loose, 2013). Furthermore, we hypothesized that decided and familiar customers have the shortest first phase, and that decided customers have a shorter second phase compared to undecided customers.

Methods

All data used in this study was collected during another research project. Therefore, in this section we describe the methods of this other research project, except for the design and analyses, which are unique for the present study.

Participants

There were 107 participants, all acquired by a convenience sample in a local supermarket. Due to incorrect task execution (n = 21) and failure of the eye tracker (n = 11), 32 participants were excluded. Subsequently, the final sample consisted of 75 participants (35 female, mean age 36.25, range = [18, 66]). All participants gave informed consent and reported normal or corrected-to-normal vision. All experimental procedures were approved by the local ethical committee of the Faculty of Social and Behavioral Sciences of Utrecht University.

Materials

Eye tracker

A Tobii Pro Glasses 2 was used to measure gaze behavior of the participants in 50 Hz. One point calibration was performed according to the manufacturers procedure.

Questionnaire

A questionnaire was used to gain insight in other variables related to the decision task of the experiment. The first part focused on the chosen product, regarding intention, impulse, preferences, satisfaction and healthiness. The second part targeted general snack habit and was derived from the Self Report Habit Index questionnaire of Verplanken and Orbell (2003). The third part asked about participants opinion regarding the environment of the supermarket, whereas the fourth part assessed their frequency of visits and familiarity with the supermarket. The fifth part of the questionnaire was only relevant to the design and goal of the original study. Lastly, the sixth part gathered demographic data and participants current level of hungriness. Part one, four and six were of particular interest for the current study. See Appendix 2 for the entire questionnaire.

Design

In the original experiment, participants were assigned to either one of two conditions. The experimental condition consisted of a health prime in the shopping basket to nudge participants to buy a healthy snack, and the control condition of a neutral prime. However, due to the non-effectiveness of this nudge and the aim of the present study, we treated all participants equally on this regard. In other words, we abandoned the original experimental condition and constructed new groups based on results from the questionnaire.

Specifically, we classified participants as decided (n = 19) or undecided (n = 56) regarding their choice with question 1 of section 1 "*Did you intend to buy this product?*". Additionally, we classified them as being familiar (n = 56) or unfamiliar (n = 19) with the supermarket with question 1 of section 4 ("*How often do* *you visit this supermarket*"). Participants were labeled as unfamiliar customers if they visited the supermarket less than once a week. This lead to a design with four groups, as visualized in Table 1.

Table 1

The amount of participants per condition

	Familiar	Unfamiliar
Decided	n = 13	n = 6
Undecided	n = 43	n = 13

Procedure

After selection and a concise briefing, participants signed the informed consent form (see Appendix 1). Thereafter, the eye tracker was installed and calibrated. Participants where then given the following instruction: "*Buy a snack (in Dutch: 'tussendoortje') of maximum €2,50. A snack is something you can eat now or later in between your three main meals*". After participants had executed the task, they walked to the checkout line where the experimenter removed the eye tracker. Participants could now do their own groceries and were subsequently asked to fill out the questionnaire. Lastly, they received a debriefing about, among other things, the goal of the experiment (see Appendix 3).

Analysis

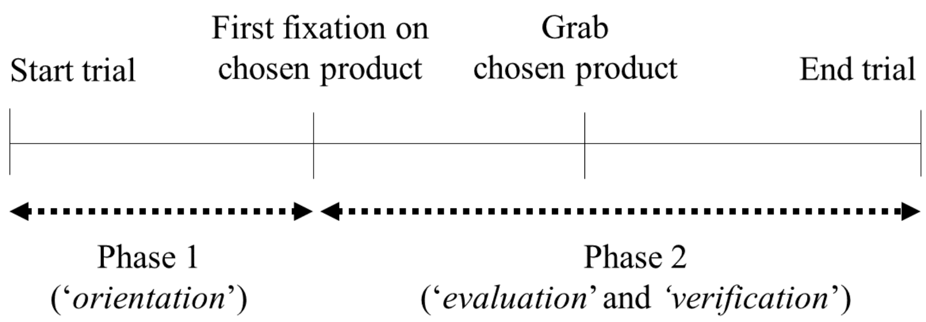
Preprocessing

First, raw eye tracking data was preprocessed and coded with GazeCode (Benjamins et al., 2018) and GlassesViewer (Niehorster et al., 2020), using the eye movements classification algorithm of Hessels et al. (2020) to determine fast phases. Due to the experimental procedure, eye tracker recordings contained large events without relevant eye tracking data. Therefore, all data before entering the supermarket (last saccade in which the entrance gates of the supermarket could be seen) and after approaching the checkout line (last saccade in which a food product could be seen) were cut-off. All saccades with data loss, often caused by blinks, were removed. We calculated the mean and standard deviation of saccadic amplitude, duration and velocity per participant, using all remaining data points.

Second, raw eye tracking data was preprocessed and coded again with GazeCode (Benjamins et al., 2018) and GlassesViewer (Niehorster et al., 2020), this time using the eye movements classification algorithm of Hessels et al. (2020) to determine slow phases. After excluding the same events as described for the fast phases, the moment in which the participant fixated the chosen product for the first time was labeled. This moment was used during subsequent analysis to distinguish the first phase from the second phase (Figure 1). Notice that this is a similar first phase cut-off as the Natural Decision Segmentation Model (Gidlöf et al., 2013), but that we did not distinguish between a second and third phase in this study. Usually these phases are separated at the last fixation at the chosen product. We decided not to do this because many participants fixated the chosen product just before the end of the trial, leaving not enough data points for the third phase to be statistically relevant. Further statistical analysis of the eye tracking and questionnaire data was done in Python 3.7, using Spyder (Anaconda3).

Figure 1

Definition of phase 1 and phase 2 in this study



Note. Notice that the first phase cut-off is similar to the Natural Decision Segmentation Model, but that we did not distinguish between a second and third phase in this study.

Statistics

(Un)decided vs. (Un)familiar. We used Shapiro-Wilk tests to test the normality of the data. For all general outcome measures at least one of the groups was significant, suggesting that the data was not normally distributed. Therefore, to test if the mean and standard deviation of saccadic amplitude, duration and velocity varied for decided and undecided participants who were either familiar or not with the supermarket, we conducted Kruskal-Wallis H-tests. The critical p-value was corrected to 0.05 / 6 = 0.0083 using Bonferroni correction because the general outcome measures were not fully independent of each other. For significant results, we conducted pairwise comparisons with Bonferroni corrected p-values (0.05 / 4 = 0.0125). We calculated Pearson's correlation coefficient to indicate effect sizes if a main effect or specific effect was significant after Bonferroni correction.

Natural Decision Segmentation Model. Next, we performed exploratory analyses to define phases in decision making using general outcome measures. We assigned all fixations and saccades prior to the first fixation at the chosen product to the first phase, and all others to the second phase. However, for some participants there was data loss regarding one or multiple saccades just prior or after the first fixation (mainly due to blinks). It was subsequently not possible to link these timestamps in the fast phases data (saccades) to the slow phases data (fixations). For these participants, the saccade closest to the saccade that actually just preceded the first fixation on the chosen product was selected (either before or after the first fixation). The maximum difference here was three timestamps (i.e. 60 ms).

To test if the duration of phase one and two were significantly different, we performed an independent t-test. We used Kruskal-Wallis H-tests to investigate which group potentially caused these difference and if these groups differed in their time to actually grab the chosen product.

Next, to analyze if the first phase was significantly different from the second phase regarding mean and standard deviation of saccadic amplitude, duration and velocity, we conducted independent t-tests with Bonferroni correction. If significant, pairwise comparison was done with Bonferroni corrected p-values. We calculated Pearson's correlation coefficient to indicate effect sizes if a main effect or specific effect was significant after Bonferroni correction. Notice that we seem to use the independent t-test and Kruskal-Wallis H-test interchangeably. Wherever we used independent t-tests, we did this because the effect size measures of the Kruskal-Wallis H-tests showed impossible values for unknown reasons. On the other hand, when performing independent t-tests, the Pearson's correlation coefficients had reasonable values. Since the significance values of the independent t-tests and the Kruskal-Wallis H-tests showed no meaningful differences, we concluded that in these cases an independent t-test was a valid way to statistically analyze this data.

Results

Data quality

Measured data loss was exceptionally high (on average 60% to 80%), but this was caused by the experimental set-up. After the eye tracker was removed from the participants head, the recording proceeded for a long time but without measuring the eye. This did not influence data analysis because all data before participants entered the supermarket and after they approached the check-out counter was excluded.

Median RMS-S2S was calculated for each participant using a 300 ms moving window. Based on these values and qualitative observations of gaze replays in GlassesViewer, we concluded that data quality was sufficient for the aim of this study. Therefore, no participants were excluded. Table 2 summarizes the mean, standard deviation and range for these medians.

Table 2

Median RMS-S2S as indicator for data quality

	Left azimuth	Left elevation	Right azimuth	Right elevation	Gaze point video x	Gaze point video y
M	1.21	0.78	1.22	0.80	31.59	22.77
SD	0.30	0.20	0.32	0.21	9.32	5.82
Range	[0.55, 1.98]	[0.37, 1.35]	[0.97, 1.98]	[0.37, 1.47]	[24.67, 55.63]	[18.76, 38.81]

Note. Azimuth and elevation values are RMS deviations in degrees, Gaze X and Y are deviations in pixels.

(Un)decided vs. (Un)familiar

To investigate if general outcome measures can account for differences in gaze behavior for customers that already decided (D) what to buy or were undecided (UD) what to buy when entering a supermarket, and who were either familiar (F) or unfamiliar (UF) with this supermarket, we analyzed (1) mean saccadic amplitude, (2) the standard deviations of saccadic amplitude, (3) mean saccadic duration, (4) the standard deviation of saccadic duration, (5), mean saccadic velocity, and (6) the standard deviations of saccadic velocity.

Standard deviations for saccadic amplitude were significantly affected by decisiveness and familiarity, $H(3) = 9.81$, $p = .020$. However, after Bonferroni correction for all six tests, the standard deviations did not significantly differ anymore. This result should therefore be interpreted with caution. Pairwise comparison with adjusted p-values showed that there were no significant differences between the groups. Given that the overall effect was not significant after correction and that the groups did not significantly differ from each other, we conclude that being (un)decided or (un)familiar did not influence the standard deviations for saccadic amplitude.

Standard deviations for saccadic velocity were significantly affected by decisiveness and familiarity, $H(3) = 8.59$, $p = .035$. However, similar to the standard deviations of saccadic amplitude, after Bonferroni correction for all six tests, the standard deviations did not significantly differ anymore. This result should therefore be interpreted with caution. Pairwise comparison with adjusted p-values showed again that there were no significant differences between the groups. Therefore, we conclude that being (un)decided or (un)familiar did not influence the standard deviations for saccadic velocity.

Mean saccadic amplitude, mean saccadic duration, the standard deviation of saccadic duration, and mean saccadic velocity were not significantly affected by being (un)decided or (un)familiar. Table 3 summarizes these results.

Table 3

General outcome measures per group

	DF	DUF	UDF	UDUF	Statistics	
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*H*	*p*
ampM	899.10 (188.45)	910.86 (163.52)	817.59 (200.31)	785.67 (172.48)	3.81	.282
ampSD	291.58 (57.46)	299.44 (48.08)	248.92 (61.00)	264.91 (33.44)	9.81	.020*
durM	63.65 (10.44)	72.79 (18.69)	66.54 (7.51)	65.23 (6.22)	2.22	.528
durSD	58.85 (11.59)	70.41 (24.50)	65.17 (13.08)	60.62 (13.82)	2.01	.406
velM	23.18 (6.55)	22.07 (5.02)	20.44 (5.52)	18.70 (5.02)	3.65	.301
velSD	16.55 (4.12)	16.79 (2.49)	14.15 (3.49)	13.26 (3.25)	8.59	.035*

Note. ampM = mean saccadic amplitude in pixels. ampSD = standard deviation of saccadic amplitude in pixels. durM = mean saccadic duration in ms. durSD = standard deviation of saccadic duration in ms. velM = mean saccadic velocity in pixels/ms. velSD = standard deviation of saccadic velocity in pixels/ms. DF = decided familiar. DUF = decided unfamiliar. UDF = undecided familiar. UDUF = undecided unfamiliar

* $p < .05$

Natural Decision Segmentation Model

Based on the Natural Decision Segmentation Model (Gidlöf et al., 2013), two phases were distinguished. The saccades and fixations prior to the first fixation at the chosen product were labeled as the first 'orientation' phase, whereas the saccades and fixations thereafter where labeled as the second 'verification and evaluation' phase.

The duration of the first phase (M = 56.64, SD = 40.89) was significantly longer than the second phase (M = 38.04, SD = 45.09), $t(3)$ = 2.65, p = .009, with a medium effect size of r = 0.30. This finding indicates that, in general, participants were in the supermarket for a longer period of time before they fixated their chosen product first compared to thereafter. However, first and second phase duration did not significantly differ between the groups. Similarly, the average duration after which a participant actually grabbed the product was not significantly affected by being (un)decided or (un)familiar. Table 4 summarizes these results.

Table 4

Phase duration and time to grab the chosen product per group

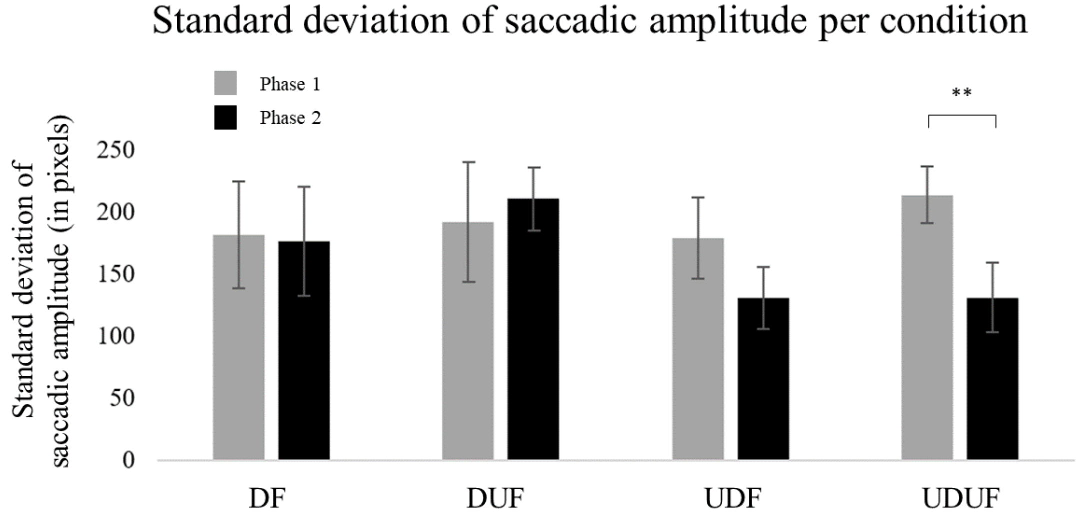
	DF	DUF	UDF	UDUF	Statistics	
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*H*	*p*
Phase 1	50.45 (33.49)	45.30 (19.46)	58.63 (46.92)	61.49 (34.36)	1.05	.790
Phase 2	43.60 (39.95)	45.40 (22.68)	37.26 (54.45)	31.70 (15.61)	3.12	.790
Grab	78.31 (48.32)	69.00 (36.76)	80.04 (63.51)	74.61 (39.69)	.08	.995

Note. All values are in seconds (s). DF = decided familiar. DUF = decided unfamiliar. UDF = undecided familiar. UDUF = undecided unfamiliar

The standard deviation of saccadic amplitude was significantly higher in the first phase compared to the second phase, also after Bonferroni correction, $t(73) = 3.86$, $p < .001$, with a medium effect size of $r = .41$. This difference can mainly be attributed to participants who were undecided unfamiliar, $t(3) = 4.12$, $p < .001$, with a large effect size of $r = .73$. This result indicates that participants who did not yet decided what to buy and were unfamiliar with the supermarket had a larger variety in the distances of their saccade before they fixated their chosen product first compared to thereafter. Figure 2 visualizes this result.

Figure 2

Standard deviation of saccadic amplitude per condition



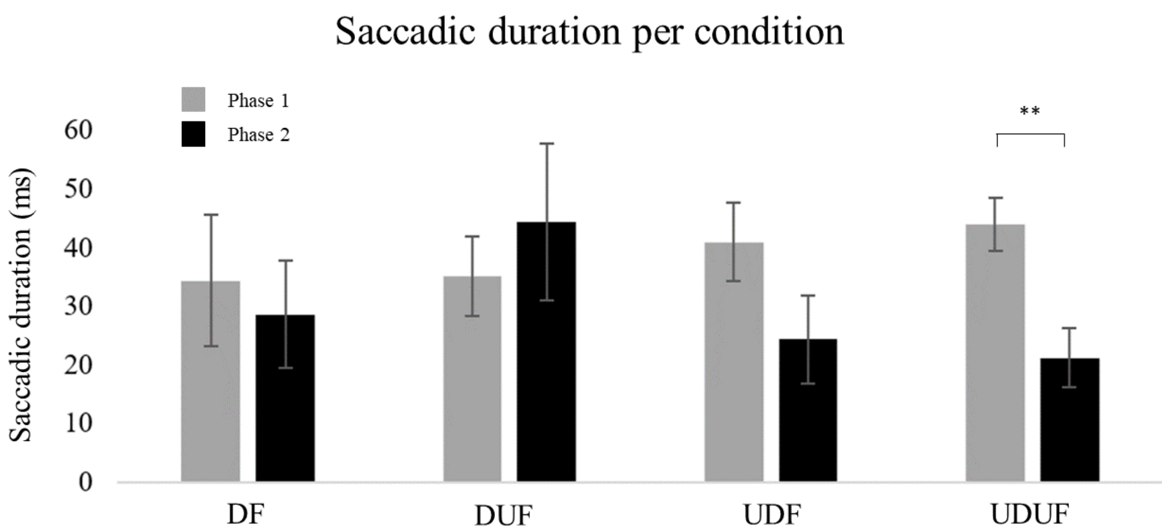
Note. DF = decided familiar. DUF = decided unfamiliar. UDF = undecided familiar. UDUF = undecided unfamiliar

** $p < .001$

Additionally, mean saccadic duration was significantly longer in the first phase compared to the second phase, also after Bonferroni correction, $t(73) = 5.31$, $p < .001$, with a large effect size of $r = .53$. This difference was mainly caused by participants who were undecided unfamiliar, $t(3) = 6.05$, $p < .001$, with a large effect size of $r = .76$. This result indicates that the saccades of participants who did not yet decided what to buy and were unfamiliar with the supermarket had a longer duration before they fixated their chosen product first compared to thereafter. Figure 3 visualizes this result.

Figure 3

Saccadic duration per condition



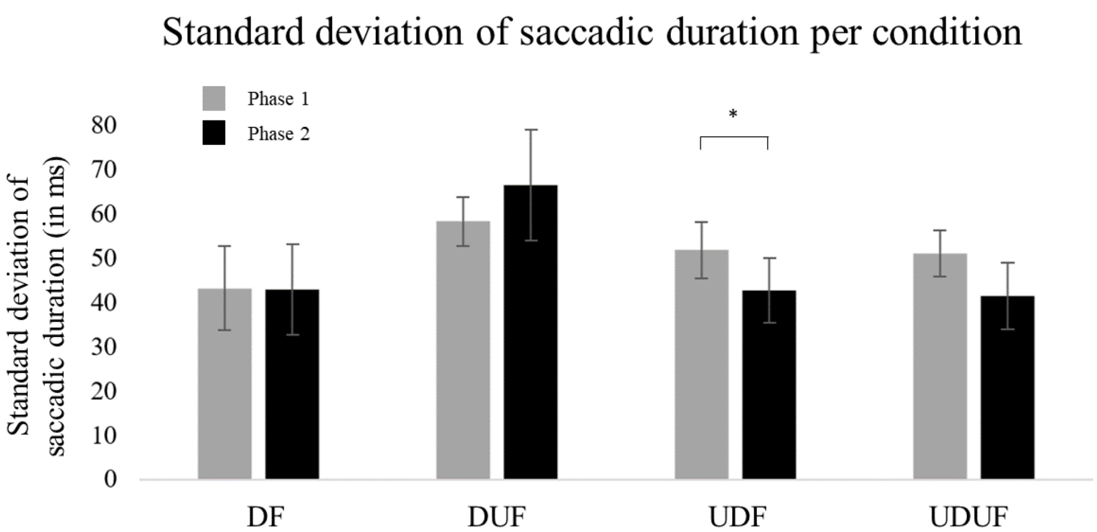
Note. DF = decided familiar. DUF = decided unfamiliar. UDF = undecided familiar. UDUF = undecided unfamiliar

** $p < .001$

Furthermore, the standard deviation of saccadic duration was higher in the first phase compared to the second phase, however, not significantly after Bonferroni correction $t(73) = 2.45$, $p = .016$. This result should therefore be interpreted with nuance. This difference was mainly caused by participants who were undecided familiar, $t(3) = 2.30$, $p = .026$. This result suggests that the saccades of participants who did not yet decided what to buy and were familiar with the supermarket had a larger variety in their duration before they fixated their chosen product first compared to thereafter. Figure 4 visualizes this result.

Figure 4

Standard deviation of saccadic duration per condition



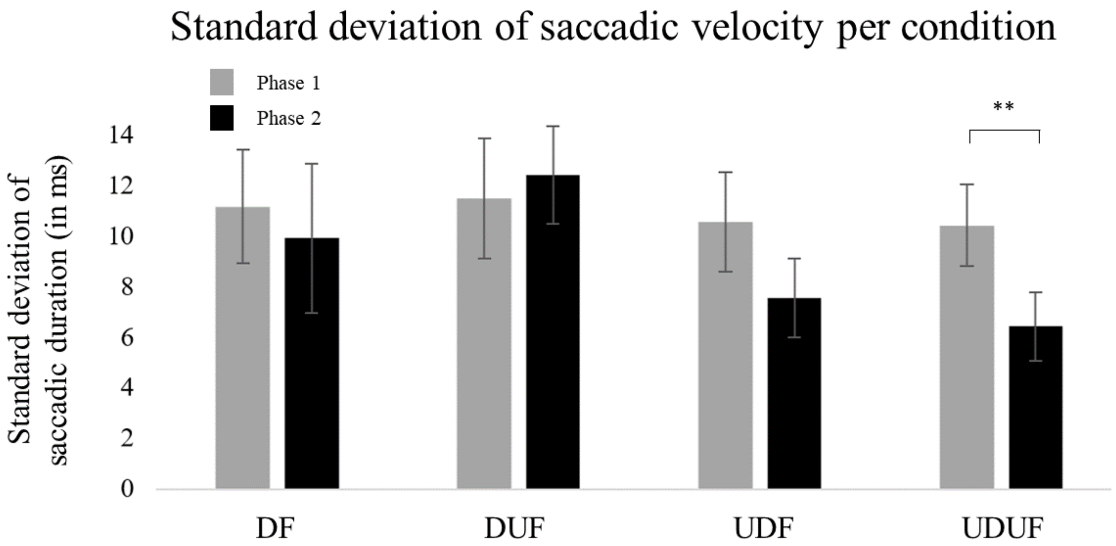
Note. DF = decided familiar. DUF = decided unfamiliar. UDF = undecided familiar. UDUF = undecided unfamiliar

* $p < .05$

Lastly, the standard deviation of saccadic velocity was also higher in the first phase, also after Bonferroni correction, $t(73) = 3.93$, $p < .001$, with a medium effect size of $r = .42$. Again, this difference stems from participants who were undecided unfamiliar, $t(3) = 3.41$, $p = .0023$, with a large effect size of $r = .56$. This result indicates that participants who did not yet decided what to buy and were unfamiliar with the supermarket had a larger variety in the speed of their saccade before they fixated their chosen product first compared to thereafter. Figure 5 visualizes this result.

Figure 5

Standard deviation of saccadic velocity per condition



Note. DF = decided familiar. DUF = decided unfamiliar. UDF = undecided familiar. UDUF = undecided unfamiliar

** $p < .001$

There were no significant differences between both phases for mean saccadic amplitude ($t(73) = 0.12$, $p = .904$) and mean saccadic velocity ($t(73) = 0.08$, $p = .935$) after Bonferroni correction. Table 5 provides an overview of these results.

Table 5

Descriptives and statistics of general outcome measures for phase 1 and 2

	Phase 1	Phase 2	Statistics		
	M (SD)	*M (SD)*	*t*	*p*	*r*
ampM	818.42 (209.94)	814.47 (191.06)	0.12	.904	-
ampSD	186.53 (68.86)	144.84 (63.47)	3.86	.00017**	.41
durM	39.88 (14.81)	26.16 (16.76)	5.31	.00000039**	.53
durSD	50.83 (13.83)	44.49 (17.66)	2.45	.016*	-
velM	20.47 (6.66)	20.39 (5.93)	0.08	.935	-
velSD	10.72 (3.92)	8.18 (4.01)	3.93	.00013**	.42

Note. Effect sizes (r) are only included when a main effect or specific effect was significant after Bonferroni correction. ampM = mean saccadic amplitude in pixels, ampSD = standard deviations of saccadic amplitude in pixels, durM = mean saccadic duration in ms, durSD = standard deviations of saccadic duration in ms, velM = mean saccadic velocity in pixels/ms, and velSD = standard deviations of saccadic velocity in pixels/ms.

* $p < .05$ ** $p < .001$

Discussion

In many cases the eyes have an essential role when it comes to making food choices. Therefore, eye tracking is a popular method to investigate decision making processes and consumer behavior. Because of technical advances, wearable eye tracker allows experiments to be conducted in complex environments like a supermarket. However, in several situations methodological challenges arise when AOI's are used to analyze the eye tracking data. In this study, we explored with wearable eye tracking if general outcome measures that do not assign AOI's can account for differences in gaze behavior for customers that already decided what to buy (resembling a search task) or did not yet decided what to buy (resembling a decision task). Additionally, we took into account their familiarity with the supermarket (resembling learning effects). The Natural Decision Segmentation Model (Gidlöf et al., 2013) was used to further explore the application of general outcome measures regarding phases in search and decision strategies.

(Un)decided vs. (Un)familiar

In line with Ajzen (2011) and Orquin & Loose (2013), we hypothesized that customers who already decided what to buy have a smaller average saccadic amplitude, shorter average saccadic duration, slower average saccadic velocity and lower standard deviations of saccadic amplitude, duration and velocity than customers who decided what to buy while being in the supermarket. Additionally, we hypothesized that that these effects are amplified if someone is familiar with the supermarket, compared to when someone is less so. Contrary to our expectations, participants who were either decided or undecided, and familiar or unfamiliar did not differ regarding their (1) mean saccadic amplitudes, (2) mean saccadic durations, (3) standard deviations of saccadic duration and (4) mean saccadic velocities. This implies that, when taking into account the entire period in which they visited the supermarket, their eyes moved for (1) similar distances, (2) similar lengths, and (4) with similar speeds, and that (3) the variety of speed in which their eyes moved was also similar. Even though these results or not in line with our expectations, we would like to stress that this analysis involved data of the entire trial rather than separated by phases according to the NDSM. We interpret the lack of differences here as an extra necessity to incorporate the NDSM phases in further analysis, as is stated in the next paragraph. Alternatively, these similarities could be attributed to a lack of power caused by the small sample size of decided unfamiliar customers. In this study, we assigned the participants to each condition post hoc based on their answers in the questionnaire. Future research could assign participants to a condition prior to the start of a trial, for example by telling them what to buy if there are in the decided condition and by strict participant selection criteria regarding their familiarity with the supermarket. This way more equal and larger group sizes could be created which would subsequently enhance statistical power.

Although the overall effect of participants who were either decided or undecided, and familiar or unfamiliar suggested differences in standard deviations for saccadic amplitude between these groups, we cannot conclude with certainty that these differences are of actual theoretical relevance for two reasons. First, after correcting for the amount of statistical tests, the overall effect was not significant anymore. Second, the specific pairwise comparisons showed no significant differences between these groups. For the same reasons, we also stress the necessity to cautiously interpret the significant overall effect of standard deviations for saccadic velocity. Based on (Park & Lessig, 1981; Hoyer, 1984; Orquin & Loose, 2013), we hypothesized that participants who where undecided unfamiliar would have higher standard deviations for saccadic amplitude and velocity than decided familiar participants. Our results might point into the direction that this is indeed the case, but given the aforementioned statistical nuances we do not make firm conclusions.

Taken these points into consideration, we cannot conclude that these general outcome measures for wearable eye tracking can account for differences for decided or undecided and familiar or unfamiliar customers in a supermarket. However, so far only the analyses over the entire length of the trials have been discussed. Given previous research on attention phases (Russo & Leclerc 1994; Gidlöf et al., 2013), it could well be that general outcome measures are still of value when taking these phases into account.

Natural Decision Segmentation Model

In general, the duration of the first phase (defined as all fixations and saccades before the first fixation on the chosen product, often labeled as 'orientation') was longer than the second phase (all fixations and saccades after the first fixation on the chosen product, often labeled as 'evaluation' and 'verification'). This could suggest that participants were generally decisive regarding their choice, however due to the experimental set-up of this study it is not possible to derive any firm theoretical conclusions from this result. For example, when people are prone to apply a maximizing strategy, they are more likely to make an objectively 'good' choice but will be less satisfied with the outcome (Schwartz et al., 2002). In general their choice process will relatively take a lot of time, thereby potentially increasing second phase duration ('comparison'). On the other hand, people who tend to apply a satisficing strategy (choice outcome is objectively worse, but they will appreciate the outcome more, Schwartz et al., 2002) would have a shorter second phase duration. Other research could take decisiveness and search strategies such as maximizing and satisficing into account to further study this relation, for example by including it in a follow-up questionnaire. Here, we simply attribute the longer duration of the first phase to the relatively short distance to the check-out counter for most chosen products, naturally decreasing the length of the second phase.

The length of first and second phase duration was not different for decided or undecided, and familiar or unfamiliar participants, whereas we expected that decided familiar participants would have had the shortest first phase duration, and undecided unfamiliar the longest. Similarly, the average duration after which a participant actually grabbed the product did not differ for decided or undecided, and familiar or unfamiliar participants. Reminiscent to general phase duration, it could again be that this result is mediated by other factors such as decisiveness (Rassin, 2007). A decided familiar customer with a tendency to be indecisive could have a longer first phase duration than an undecided unfamiliar customer with a tendency to be decisive. (Patalano, Juhasz, & Dicke, 2009) Although we currently did not find a difference between these groups, we believe it is fruitful to investigate a possible relation further in future research for the following reason. It is interesting to note that first phase duration somewhat resembles T_{50}, a measure used to establish how long it takes for 50% of a population to fixate a target area for the first time, often a brand logo in an advertisement (Hooge & Camps, 2013). In the context of a supermarket and from the perspective of the retailer, the chosen product can be defined as the target area. In this study, participants had to buy a snack, and the original set-up of the experiment was focused at nudging participants to buy a healthy snack. First phase duration could, similarly to T_{50}, indicate the attention attracting power of such a healthy snack, and provide an indication to the retailer how visible their product is. This general outcome measure could be of particular value when undecided customers would have had longer first phase durations. The retailer could benefit most from making their product visible to this group of people sooner, thereby increasing the likelihood for this product to enter the consideration set early on and thereafter being chosen. To explore this idea further, we calculated T_{50} (in seconds) in a follow-up analysis: decided familiar customers = 44.53; decided unfamiliar = 50.72; undecided familiar = 51.99; undecided unfamiliar = 49.93. As expected, decided familiar participants seemed to be quickest in fixating their chosen product first, but we did not conduct any statistical analysis to further substantiate this explorative follow-up analysis. Future research could further investigate the application of T_{50} in situations similar to this study.

The standard deviations of saccadic amplitude and velocity, and mean saccadic duration were different in the first phase compared to the second. For all three of these general outcome measures the difference can be attributed to the participants who were undecided unfamiliar because they had higher standard deviations of saccadic (1) amplitude and (2) velocity, and longer (3) mean saccadic durations. Similarly, (4) the standard deviations of saccadic duration were higher in the first phase, but this difference arose from participants who were undecided familiar. These results are in line with our expectations regarding complete trial duration (instead of separated for two phases), and might explain the lack of differences found when analyzing these general outcome measures for the entire trial. Based on our results, it seems reasonable to conclude that customers who did not yet decided what to buy when entering a supermarket and are unfamiliar with this supermarket, are more likely to show viewing patterns with more variety in the (1) distance and (2) speed of their saccades before they fixate the chosen product for the first time than customers who either already know what they want to buy, are familiar with the supermarket, or both. Additionally, (3) the duration of their saccades is longer in this same phase. Previous research defined this first phase as 'orientation'. More variety in saccadic distance and speed, and longer saccadic durations are plausible attributes of such an orientation phase, given that customers want or need to explore their environment and that it is therefore beneficial to scan a large proportion of the visual field. This is especially the case for undecided unfamiliar customers. Whereas other studies using the NDSM showed a distinct pattern for the first phase regarding fixations and re-dwells while using AOI's (Gidlöf et al., 2013), our results suggest that general outcome measures also contribute to the conceptual differences between these phases.

However, we also want to stress that there are factors which make this relation less straightforward. First, it is not just customers who are undecided unfamiliar who give rise to differences in general outcome measures regarding the first 'orientation' phase, but also (4) customers who were undecided familiar and had higher standard deviations of saccadic duration in this first phase. Second, this study found no differences in mean saccadic amplitude and velocity. Third, the general outcome measures of amplitude and velocity are tightly connected with velocity, given that the latter is a result of the former two. Fourth, usually a third phase is defined (all saccades and fixations after the last fixation on the chosen product) but we did not do so because of the limited amount of data points for a potential third phase in this data set. This might have created a flaw when comparing the first phase with the second (and implicit third). Given the experimental set-up and data analysis of this study we cannot disentangle these factors, but we encourage future research to do so. For example, saccadic amplitude, duration and velocity could be experimentally manipulated, participants could purposefully be assigned to one of the four groups instead of doing this post hoc, or the third phase could be included.

Limitations, suggestions and applications

Besides the aforementioned limitations of this study and suggestions for future research, we would like to point out several other key issues.

First of all, in this research we defined being familiar with the supermarket as visiting this supermarket at least once a week. All participants who indicated that they visited this supermarket less than once a week were labeled as unfamiliar. It can be argued that this criterium has low ecological validity, thus not being completely representative of someone's familiarity with a supermarket. Additionally, with our current criterium only 19 participants were labeled as unfamiliar compared to 56 familiar participants, creating unequal group sizes. Indeed, we suggest for future research to improve the selection methods regarding the familiarity of participants, for example by asking specific questions about the supermarket (for instance: "*Could you point out where* x *is located?*").

Second, in the original study participants were allowed to buy only one item, whereas it is common to buy multiple products when visiting a supermarket. This was convenient for the purpose of the present study, since the NDSM does not define phases when multiple products are bought. Indeed, it would be challenging to investigate which fixations and saccades are relevant for the purchase of which product. The NDSM and related studies are a promising start to identify gaze behavior of consumers in complex settings, but further research should consider to include the purchase of multiple products.

Third, we only used standard deviations of saccadic amplitude, duration and velocity as dispersion measures, but other studies using different measures for variability suggest that these measures could also be relevant for research on decision making processes and its phases. For example, Hooge and Camps (2013) showed how entropy, an indication of the uncertainty of a variable's possible outcomes, can show the ability of advertisements to guide gaze patterns. The finding that undecided unfamiliar customers have larger standard deviations in their saccadic amplitude and velocity during the first 'orientation' phase could be substantiated by using other dispersion methods such as entropy. Note that this application of entropy involves the use of scan paths and therefore also AOI's. Alternatively, variability could also be operationalized using walking patterns or head movements that direct gaze.

Fourth, we suggest to complement this study using general fixation measures such as fixation duration. For instance, Glöckner and Herbold (2011) found that the first 10 to 20 fixations in a choice task have a relatively large proportion of short fixations, so it would be interesting to investigate if this same observation holds when using the NDSM or studying decided, undecided, familiar and unfamiliar customers.

In the introduction we argued that AOI's are difficult to implement in measurements made in complex environments like a supermarket. First, it can be problematic to define meaningful AOI's due to the amount of relevant features in a supermarket or overlapping AOI's when participants need to choose a product. Second, it can be challenging to reliably assign fixations to AOI's because of product placement and size, or eye tracker precision. Third, it can be difficult to interpret the results of the analyses since attention is not necessarily located at the point of fixation, gaze behavior can be related to other events such as physical activity instead of cognitive processes, and various cognitive processes could be at play. The general outcome measures used in this study can counteract the first and second issue because no AOI's have to be defined and no fixations have to be assigned to AOI's. Therefore, when these challenges arise in future eye tracking research using AOI's, we suggest to use general outcome measures instead. If this is not possible for that specific research, we advocate to perceive these challenges with AOI's as a signal detection problem. When an AOI is enlarged to include all actual fixations at this AOI, it is also more likely to include fixations that belong to nearby features of the environment (i.e. false alarms). Contrary, when it is made smaller to decrease the amount of false alarms, it will naturally also decrease the amount of hits. We refer to Orquin et al. (2016) and Hessels et al. (2016) for guidelines on how to deal with this issue specifically. Furthermore, when creating software for manual classification and data quality checks like Gazecode (Benjamins et al., 2018) and GlassesViewer (Niehorster et al., 2019), we propose to vary the size of the gaze cursor based on the precision of the eye tracker, or even for a specific part of a trial. This could substantially support the reliability of assigning fixations to AOI's. Lastly, when poor data quality challenges the ability to assign fixations to AOI's, it should be considered if data of one or two eyes are used in these analyses (Hooge, Holleman, Haukes, & Hessels, 2018).

The third challenge of interpreting the results of AOI based analysis, however, still stands when using general outcome measures. We assume that the higher variability of saccadic distance and speed for undecided unfamiliar customers in the first 'orientation' phase also means that the allocation of their attention was more dispersed (overt attention), but since the location of fixating and attending can be decoupled (covert attention), it stays challenging to interpret this result. Moreover, general outcome measures also do not indicate whether at a particular moment gaze behavior reflected cognitive processes or physical movements such as maintaining distance to other customers or grabbing a product. Lastly, even if it would be known that general outcome measures reflect cognitive processes at a particular point, it still stays a challenge to know which cognitive process is at play. This is especially difficult when multiple products can be purchased instead of just one, as pointed out earlier. We value the awareness that these challenges highlight that an eye tracker merely indicates where someone is looking and that everything else is a deduction from this, but also appreciate the practical value eye tracking can have in various situations. It remains a topic for future research to improve interpretations of eye tracking results, also when general outcome measures are used to counteract other AOI related challenges.

Conclusion

Gaze behavior is key when making food decisions in a supermarket. However, we identified three issues when using wearable eye tracking and AOI's to investigate these processes in complex environments and postulated that general outcome measures can overcome two of these challenges. We showed that customers who did not yet decided what to buy before they entered the supermarket and who are unfamiliar with this supermarket have more variety in the distance and speed of their saccades before they fixate the chosen product first compared to other customers. Additionally, on average their eyes moved for longer durations in this same first phase. These findings substantiate the NDSM by providing more evidence for an orientation strategy, taking place before the product that is eventually chosen is fixated for the first time. Although general outcome measures seem to be a promising alternative when AOI based analysis is unfeasible, there are still many topics to address before we fully understand the relation between gaze behavior and making food decisions in a complex environment.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179-211.

Benjamins, J. S., Hessels, R. S., & Hooge, I. T. (2018). Gazecode: open-source software for manual mapping of mobile eye-tracking data. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications* (pp. 1-4).

Danner, L., de Antoni, N., Gere, A., Sipos, L., Kovács, S., & Dürrschmid, K. (2016). Make a choice! Visual attention and choice behaviour in multialternative food choice situations. *Acta Alimentaria*, *45*(4), 515-524.

Gidlöf, K., Anikin, A., Lingonblad, M., & Wallin, A. (2017). Looking is buying. How visual attention and choice are affected by consumer preferences and properties of the supermarket shelf. *Appetite*, *116*, 29-38.

Gidlöf, K., Wallin, A., Dewhurst, R., & Holmqvist, K. (2013). Using eye tracking to trace a cognitive process: Gaze behaviour during decision making in a natural environment.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, *62*, 451-482.

Glaholt, M. G., Wu, M. C., & Reingold, E. M. (2010). Evidence for top-down control of eye movements during visual decision making. *Journal of vision*, *10*(5), 15-15.

Glöckner, A., & Herbold, A. K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*, *24*(1), 71-98.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in cognitive sciences*, *9*(4), 188-194.

Hessels, R. S., Kemner, C., van den Boomen, C., & Hooge, I. T. (2016). The area-of-interest problem in eyetracking research: A noise-robust solution for face and sparse stimuli. *Behavior research methods*, *48*(4), 1694-1712.

Hessels, R. S., Niehorster, D. C., Holleman, G. A., Benjamins, J. S., & Hooge, I. T. (2020). Wearable technology for 'real world reserach' – realistic or not? Manuscript submitted for publication.

Hessels, R. S., Niehorster, D. C., Nyström, M., Andersson, R., & Hooge, I. T. (2018). Is the eye-movement field confused about fixations and saccades? A survey among 124 researchers. *Royal Society open science*, *5*(8), 180502.

Hessels, R. S., van Doorn, A. J., Benjamins, J. S., Holleman, G. A., & Hooge, I. T. (2020). Task-related gaze control in human crowd navigation. *Attention, Perception, & Psychophysics*, 1-20.

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

Holmqvist, K., Nyström, M., & Mulvey, F. (2012, March). Eye tracker data quality: what it is and how to measure it. In *Proceedings of the symposium on eye tracking research and applications* (pp. 45-52).

Hooge, I. T., & Camps, G. (2013). Scan path entropy and arrow plots: Capturing scanning behavior of multiple observers. *Frontiers in Psychology*, *4*, 996.

Hooge, I. T., Holleman, G. A., Haukes, N. C., & Hessels, R. S. (2019). Gaze tracking accuracy in humans: One eye is sometimes better than two. *Behavior Research Methods*, *51*(6), 2712-2721.

Hoyer, W. D. (1984). An examination of consumer decision making for a common repeat purchase product. *Journal of consumer research*, *11*(3), 822-829.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence*, *20*(11), 1254-1259.

Johansson, R. S., Westling, G., Bäckström, A., & Flanagan, J. R. (2001). Eye–hand coordination in object manipulation. *Journal of Neuroscience*, *21*(17), 6917-6932.

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, *108*(33), 13852-13857.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities?. *Vision research*, *41*(25-26), 3559-3565.

Lohse, G. L. (1997). Consumer eye movement patterns on yellow pages advertising. *Journal of Advertising*, *26*(1), 61-73.

March, J. G. (1978). Bounded rationality, ambiguity, and the engineering of choice. *The Bell Journal of Economics*, 587-608.

Milosavljevic, M., Navalpakkam, V., Koch, C., & Rangel, A. (2012). Relative visual saliency differences induce sizable bias in consumer choice. *Journal of Consumer Psychology*, *22*(1), 67-74.

Niehorster, D. C., Hessels, R. S., & Benjamins, J. S. (2020). GlassesViewer: Open-source software for viewing and analyzing data from the Tobii Pro Glasses 2 eye tracker. *Behavior Research Methods*, 1-10.

Orquin, J. L., Ashby, N. J., & Clarke, A. D. (2016). Areas of interest as a signal detection problem in behavioral eye-tracking research. *Journal of Behavioral Decision Making*, *29*(2-3), 103-115.

Orquin, J. L., & Holmqvist, K. (2018). Threats to the validity of eye-movement research in psychology. *Behavior research methods*, *50*(4), 1645-1656.

Orquin, J. L., & Loose, S. M. (2013). Attention and choice: A review on eye movements in decision making. *Acta psychologica*, *144*(1), 190-206.

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

Over, E. A. B., Hooge, I. T. C., Vlaskamp, B. N. S., & Erkelens, C. J. (2007). Coarse-to-fine eye movement strategy in visual search. *Vision Research*, *47*(17), 2272-2280.

Park, C. W., & Lessig, V. P. (1981). Familiarity and its impact on consumer decision biases and heuristics. *Journal of consumer research*, *8*(2), 223-230.

Patalano, A. L., Juhasz, B. J., & Dicke, J. (2010). The relationship between indecisiveness and eye movement patterns in a decision making informational search task. *Journal of Behavioral Decision Making*, *23*(4), 353-368.

Peschel, A. O., & Orquin, J. L. (2013). A review of the findings and theories on surface size effects on visual attention. *Frontiers in psychology*, *4*, 902.

Posner, M. I. (1980). Orientation of attention. *Quarterly journal of experimental psychology*, *32*(1), 3-25.

Rassin, E. (2007). A psychological theory of indecisiveness. *Netherlands journal of psychology*, *63*(1), 1-11.

Russo, J. E., & Leclerc, F. (1994). An eye-fixation analysis of choice processes for consumer nondurables. *Journal of consumer research*, *21*(2), 274-290.

Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: Happiness is a matter of choice. *Journal of personality and social psychology*, *83*(5), 1178.

Shepherd, M., Findlay, J. M., & Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology*, *38*(3), 475-491.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, *69*(1), 99-118.

van der Laan, L. N., Hooge, I. T., De Ridder, D. T., Viergever, M. A., & Smeets, P. A. (2015). Do you like what you see? The role of first fixation and total fixation duration in consumer choice. *Food Quality and Preference*, *39*, 46-55.

Van der Stigchel, S., Belopolsky, A. V., Peters, J. C., Wijnen, J. G., Meeter, M., & Theeuwes, J. (2009). The limits of top-down control of visual attention. *Acta psychologica*, *132*(3), 201-212.

Van Herpen, E., & Van Trijp, H. C. (2011). Front-of-pack nutrition labels. Their effect on attention and choices when consumers have varying goals and time constraints. *Appetite*, *57*(1), 148-160.

Van Loo, E. J., Nayga Jr, R. M., Campbell, D., Seo, H. S., & Verbeke, W. (2018). Using eye tracking to account for attribute non-attendance in choice experiments. *European Review of Agricultural Economics*, *45*(3), 333-365.

Verplanken, B., & Orbell, S. (2003). Reflections on past behavior: a self-report index of habit strength 1. *Journal of applied social psychology*, *33*(6), 1313-1330.

Yarbus, A. L. (1967). Eye movements during perception of complex objects. In *Eye movements and vision* (pp. 171-211). Springer, Boston, MA.

*[**NOTE**: This informed consent was made during the original study, **not** during the thesis of Tiemen Wagenvoort]*

APPENDIX 1: Informed consent

Toestemmingsverklaring voor deelname aan wetenschappelijk onderzoek "Looking at the role of attention in nudge effectiveness"

Beste deelnemer,

Dit formulier is bedoeld om u te informeren over de werkwijze van het onderzoek bij Sociale, Gezondheid en Organisatie Psychologie, Universiteit Utrecht, en om uw toestemming te vragen voor deelname.

Het doel van deze studie is om beter te begrijpen hoe mensen beslissingen maken in de supermarkt. Daarom gaat u zo boodschappen doen terwijl u een eye-tracker draagt die uw oogbewegingen vastlegt. Daarna zullen er nog een aantal vragen volgen over o.a. hoe u meestal beslissingen neemt en uw sociodemografische gegevens. Het totale onderzoek zal ongeveer 15 minuten in beslag nemen.

Deelname is geheel vrijwillig en u kunt op elk moment stoppen, zonder dat u hiervoor een reden hoeft te geven. Er zijn geen goede of foute antwoorden. Wees in uw reacties alstublieft zo eerlijk mogelijk door uw eigen persoonlijke voorkeur aan te geven. De gegevens die worden verzameld zullen vertrouwelijk worden behandeld en zullen voor onderzoeksdoeleinden minimaal 10 jaar bewaard blijven. Er zullen geen namen of andere persoonlijke informatie over uw identiteit worden gevraagd, gebruikt of opgeslagen. Op de beelden van de eye-tracker zal uw oog te zien zijn en mogelijk delen van uw lichaam zoals uw armen of benen op het moment dat u daarnaar kijkt. Echter, deze beelden zullen gecodeerd worden opgeslagen en zullen dus niet direct herleidbaar zijn naar uzelf.

Als u vragen heeft kunt u contact opnemen met de onderzoeker door te mailen naar: f.e.deboer@uu.nl

Door hieronder uw handtekening te plaatsen geeft u aan akkoord te gaan met het volgende:

Ik ben geïnformeerd over het onderzoek. Ik heb de schriftelijke informatie gelezen. Ik heb de mogelijkheid gekregen om vragen te stellen over het onderzoek. Ik heb gelegenheid gekregen om over mijn deelname aan het onderzoek na te denken en die is geheel vrijwillig. Ik heb het recht om te allen tijde de toestemming die ik geef weer in te trekken en mijn deelname aan het onderzoek stop te zetten zonder opgaaf van redenen.

Handtekening: Datum:

*[**NOTE**: This questionnaire was made during the original study, **not** during the thesis of Tiemen Wagenvoort]*

APPENDIX 2: Questionnaire

Universiteit Utrecht

Datum: ………/………/…….. **Tijd:** ______:______

PP nr: ________

Vragenlijst supermarktonderzoek

1. De volgende vragen gaan over het tussendoortje dat u hebt gekozen.

Welk tussendoortje heeft u gekozen?

………

Omcirkel het antwoord dat het beste met uw mening overeenkomt.

Ik was van plan dit product te kopen.	Ja	Nee
Ik koop dit product vaak in de supermarkt.	Ja	Nee
Ik heb dit product in een impuls gekocht.	Ja	Nee
Ik heb een sterke voorkeur voor mijn gekozen product.	Ja	Nee
Ik ben tevreden met mijn keuze.	Ja	Nee
Ik vind dit product een gezonde keuze.	Ja	Nee

2. De volgende vragen gaan over tussendoortjes in het algemeen. Omcirkel het nummer dat het beste met uw mening overeenkomt.

Een tussendoortje kopen in de supermarkt is iets …	**Helemaal mee oneens**					**Helemaal mee eens**
wat ik vaak doe.	1	2	3	4	5	6	7
wat me moeite zou kosten om niet te doen.	1	2	3	4	5	6	7
wat hoort bij mijn routine.	1	2	3	4	5	6	7
wat typisch iets voor mij is.	1	2	3	4	5	6	7

3. Omcirkel bij de volgende stellingen welk nummer het beste met uw mening overeenkomt.

	Helemaal mee oneens						**Helemaal mee eens**
Ik vind het belangrijk om gezond te eten.	1	2	3	4	5	6	7
In mijn supermarkt zijn gezonde producten aanwezig.	1	2	3	4	5	6	7
In mijn supermarkt is het makkelijk om gezond boodschappen te doen.	1	2	3	4	5	6	7
Anderen in mijn supermarkt kopen gezonde producten.	1	2	3	4	5	6	7

4. De volgende vragen gaan over uw boodschappengedrag. Vul in wat bij u van toepassing is.

Hoe vaak doet u boodschappen in deze winkel?	❏ Vaak, ... keer per week ❏ Soms, ... keer per maand ❏ Bijna nooit, namelijk: ....
Hoe lang komt u al in deze supermarkt?	❏ Korter dan 2 maanden ❏ 2 tot 6 maanden ❏ Langer dan 6 maanden

5. De volgende vragen gaan over de afbeeldingen in het winkelmandje. Vul deze in.

Heeft u gekeken naar de afbeeldingen in het winkelmandje? *Indien antwoord 'nee', ga door naar vraag 6.*	Ja	Nee

Hoe uitgebreid hebt u naar de afbeeldingen in het winkelmandje gekeken op een schaal van 1-7?

Helemaal <u>niet</u> uitgebreid						Heel erg uitgebreid
1	2	3	4	5	6	7

Welke afbeeldingen werden er getoond in het winkelmandje?	❏ .... ❏ Ik weet het niet

6. De volgende vragen gaan over uw algemene gegevens. Vul deze in.

Hoeveel honger heeft u nu op een schaal van 1-10?

Helemaal <u>geen</u> honger									Heel erg honger
1	2	3	4	5	6	7	8	9	10

Wat is uw leeftijd?	.... jaar
Wat is uw geslacht?	❏ Man ❏ Vrouw ❏ Anders
Wat is uw lengte?	.... centimeter
Wat is gewicht?	.... kilogram
Wat is uw hoogst afgeronde opleidingsniveau?	❏ Lagere school (basisschool) ❏ LBS, vso (lts, leao, vbo, huishoudschool, ambachtsschool) ❏ VMBO/MAVO (lwoo, ulo, mulo) ❏ havo/vwo (mms, hbs) ❏ MBO (mts, meao, middenstandsdiploma, pdb, mba) ❏ HBO (hts, heao, kweekschool) ❏ Universiteit

*[**NOTE**: This debriefing was made during the original study, **not** during the thesis of Tiemen Wagenvoort]*

APPENDIX 3: Debriefing

Uitleg en nabespreking onderzoek van het wetenschappelijke onderzoek “Looking at the role of attention in nudge effectiveness”

Bedankt voor uw deelname aan het onderzoek. Het doel van dit onderzoek was om beter te begrijpen hoe mensen beslissingen nemen in de supermarkten de rol van aandacht hierin. In dit onderzoek waren er twee condities: 1) een supermarkt waar wij niets aan veranderd hadden (controle conditie) en 2) een supermarkt waarbij wij probeerde de gezonde keuze opvallender te maken doordat u een boodschappenmandje kreeg met daarin gezonde producten weergegeven.

We hebben gekeken hoe deze verandering in de supermarkt uw aandacht en aankoopgedrag heeft beïnvloed ten opzichte van de controle conditie waar niets veranderd was. Wij verwachten dat een kleine aanpassing in de supermarktomgeving de gezondere keuze meer laat opvallen en waardoor de gezonde optie vaker gekozen wordt.

Als u vragen heeft over in dit onderzoek kunt u contact opnemen met Femke de Boer via email: f.e.deboer@uu.nl

Vriendelijk verzoek om de inhoud van dit onderzoek niet te bespreken met mensen die (misschien) nog aan het onderzoek gaan meedoen. Voor het onderzoek is het van belang dat deelnemers van tevoren niet op de hoogte zijn van de onderzoeksvraag en onderzoeksmethode.

Nogmaals bedankt!