Tracking the eye of the beholder:
The roles of effort and preferences in everyday decision making

Eye-tracking i mataffären:
Hur kognitiv ansträngning och preferenser påverkar vardagligt beslutsfattande

Erik Jansson

Handledare / supervisors

Annika Wallin
Kerstin Gidlöf

KOGM20

2016-03-14

Masteruppsats i kognitionsvetenskap
Avdelningen för kognitionsvetenskap
Filosofiska institutionen
Lunds universitet

Master's thesis (2 years) in Cognitive Science
Lund University Cognitive Science
Department of Philosophy
Lund University

Tracking the eye of the beholder:
The roles of effort and preferences in everyday decision making

Erik Jansson
erik.jansson@fil.lu.se

This study investigates the relation between preferences and cognitive effort in decision making tasks in a natural setting. Does the way a supermarket's shelves are organized affect how easy it is to choose what items to purchase? Does it affect the quality of our choice?
The eye-movements of participants were recorded during their regular shopping trips in two supermarkets with different shelf organizations. They also responded to surveys about their shopping routines and product preferences.

In the analysis, the participants were divided into groups according to their stated product preferences: one group who strongly favored certain sub-categories of products (e.g. ecological pasta); one group who strongly favored certain brands and one control group who had no strong preference either way. This division meant that each of the first two groups' preferences were either congruent or incongruent with the organization of the shelves in the respective supermarket. The control group was used as a base-line.

Measurements of total time and evaluation time in front of each shelf; number of products looked at; number of returns to products already looked at and self-reported difficulty in choosing products were used as indicators of cognitive effort in finding and choosing products. Preferences were matched with product attributes to calculate the quality of purchase options and ultimate decision quality.

The results suggest that participants used less effort in choosing their products when the shelf was organized in a way congruent with their preferences (e.g. by category for a category-focused participant). Results also indicate that congruence entails lower quality purchase decisions. That is, the attributes of the chosen products match the participant's stated preferences poorly. It is argued that when cognitive resources are freed up because less effort is used in finding and choosing a product, those resources are not necessarily used to make better decisions. Rather, the less effort that is used throughout the whole process, the worse the purchase decisions will be.

1 Introduction

Supermarkets are for everyone, so they have to deal with a large variety of customer preferences. Thus there are large parts of its range of products that you probably have never even considered buying or even paid attention to. Imagine instead that there was a supermarket just for you. What would it look like? Would you like more options than the ones you usually buy? Would you still like everything sorted the usual way or would you want it organized some other way? If you could have it your way it would be so much easier to find and purchase what you need!

Supermarkets of course have to contain much more than just the things you would buy and no one person's personalized supermarket would look quite like another's. There is a trade-off between having a wide range of products on the one hand and difficulties for everyone to find what they want on the other. It is therefore incumbent upon the retailers to organize their store and its products in a way that makes it as easy as possible for the customer—regardless of preference—to find what they are looking for. At least insofar as it makes fiscal sense. Individual stores now carry as many as 100,000 different products (Ball, 2004) while 85% of a typical family's needs are met by a mere 150 products (Roberts, 2003)!

Let's say you are in the market for some fruit yoghurt. Perhaps you would only go for one of the healthier options. Most brands carry "light", low sugar and low fat options, so there is no lack of yoghurts for you to choose from. There is one problem however: in the particular supermarket you are in the yoghurts are sorted by brand and not by type. This means you have to go through each brand; find their healthier options and then memorize attributes and prices to be able to compare to other options—that is, if you can find them among their respective brands.

You may say that this is just a bad layout. But is it bad for everyone? What about someone who is brand-loyal and more open to different types of yoghurt? What is a terrible layout for you turns out to be perfect for them. They can immediately disregard everything else on the shelf and just focus on that particular brand's assortment. And what about those who want ecological options? They do not want the shelf to be organized by brand either, but since there is most likely only a partial overlap between light yoghurts and those that are ecological, the eco-friendly person would have as hard a time finding their options on your perfect shelf as you did on the brand-organized one.

Questions of supermarket assortment and consumer preference may seem of little interest to cognitive science, but there are parallels between supermarket shopping and psychological lab experiments to be found. The supermarket is in a way a great big laboratory for a scientist with the right experimental design and equipment. It is a place where visual attention is important; decision making central and where your working memory is working overtime. It is a cognitively taxing environment, especially if you find yourself in an unfamiliar supermarket. There is often time-pressure combined with the fatigue at the end of a work day.

While shopping, most of us are blissfully unaware of these complex, cognitive abilities and processes that underlie our behavior. They are however central to this paper and will be outlined below.

Decision making and cognitive effort

Oftentimes we can't wait to get out of the supermarket and the decisions we have to make are part of what's keeping us inside. What to get for dinner, what brand to get, how much to get etc. In cognitive terms, decisions can be divided into action-outcome decisions and stimulus-response decisions (Gazzaniga et al., 2014). Action-outcome decisions are made after more or less conscious evaluation of the expected result. Stimulus-response decisions are habitual and prompted by the context.

When in the supermarket we tend to make both of these types of decisions. Perhaps when buying milk, all we need is to see the one we usually buy to make the decision to buy it—a typical example of a habitual or stimulus-response decision. In another section of the store, perhaps you are buying eggs. You want them to come from free range chickens but not be too expensive. In this case you may have to spend more time comparing attributes.

When a person evaluates the expected outcome of a decision, it is in terms of reward value the comparison is made. This value is subjective and is affected by several factors. One of the more relevant to supermarket shopping is that of effort (Bettman, Luce & Payne, 1998). How much time and effort are you willing to put into finding the right eggs? Is the difference between the perfect eggs and the merely adequate eggs really worth the extra effort put in? Context is of course also important. If you are in a hurry or just very hungry, you will probably opt for the merely satisfactory option.

A seemingly simple decision such as choosing what eggs to buy makes "severe demands […] upon the choosing organism" (Simon, 1955, p. 103) because of all the possible pay-offs and all the information needed to make an optimal decision. Instead we make approximations and simplify our decision making tasks. Thus, we must take into account not only the expected outcome of a decision to be made but also the effort needed to make it.

Rushworth et al. have argued (2012) that it is also fundamental to the way the brain processes decision making that it is not only a comparison between the value of different options that is made, but that the present option is weighed against the prospect of finding a better option.

Many consumers arrive at the yoghurt shelf already determined to buy a certain kind of yoghurt as in the case of the eco-friendly and the health-conscious shoppers example. Some, however, narrow their scope only at the sight of the overwhelming assortment. "I'll just get something cheap" is as much a strategy as it is the last resort of the tired shopper. A strategy to what end you may ask? It is a strategy to minimize cognitive effort. A consumer may easily eliminate almost all options by focusing on only one or two attributes, e.g. price, a process called elimination by aspects (Tversky, 1972). After all, the amount of information a person can hold in working memory at any time is very limited (7 +/- 2 according to Miller, 1956). The required cognitive effort grows exponentially with the number of considered items when every attribute is taken into account, so there is a lot of effort to save by doing this.

In her 1994 article *Measures of cognitive effort*, Elizabeth Cooper-Martin introduced a model where the basic cognitive effort unit is a comparison between products (Cooper-Martin, 1994). A measure of cognitive effort while choosing a product can include number of attributes processed, number of alternatives processed and number of comparisons processed (Cooper-Martin, 1994; Shugan, 1980). Decision time, multiple processing (processing several attributes and/or alternatives) and self-reports have all been shown to be reliable measures of cognitive effort (Cooper-Martin, 1994; Bettman, Johnson & Payne, 1990).

Decision time as described by Cooper-Martin, is the number of seconds between the moment the subject sees all the products and the moment when she indicates her choice.

Most of our decision making effort is spent during what Russo & Leclerc (1994) call *the evaluation stage*. That is, in comparing the attributes of the products rather than merely scanning the assortment. This activates several aspects of the working memory (Baddeley & Hitch, 1974). With our focus on cognitive effort it becomes important to be able to draw a reliable line between the orientation stage (scanning) and the evaluation stage since a majority of effort is exuded during the latter. Building on the model presented by Russo & Leclerc (1994), Gidlöf et al. (2013) developed the *Natural Decision Segmentation Model (NDSM)* which established the evaluation stage's start at the point where the item that was ultimately chosen was first looked at (fig 1.).

Fig. 1. A timeline of a participant's decision stages while looking at a supermarket shelf according to the natural decision segmentation model (NDSM).
1: The first fixation (on any product) on the shelf.
2: The first fixation on the product that was ultimately chosen.
3: The last fixation on the chosen product.
4: The last fixation (on any product) on the shelf.
The red line represents the *orientation* stage between 1 and 2.
The green line represents the *evaluation* stage between 2 and 3.
The blue line indicates the *verification* stage between 3 and 4.

Using Cooper-Martin's decision time measure, decision time includes both orientation time and evaluation time. It is therefore possible to better focus on actual decision making by drawing the line at the start of the evaluation phase and leaving out time spent searching the shelves without making any comparisons between products.

The evaluation stage would still include the majority of the multiple processing since that is predominately made up of product comparisons. In our study, we will be looking at both total amount of time looking at products (*dwell* time) and evaluation stage dwell times as well as re-dwells (returning to an item previously looked at) as indicators of cognitive effort. In a natural setting such as a supermarket, self-reported measures of cognitive effort cannot be interspersed without also interrupting the shopping experience. Instead of somehow having the participants report their cognitive effort throughout the experiment, we opted to instead make it part of the post-experiment surveys. One effort measurement per product category, measuring the overall difficulty of choosing said product.

Decision making: lab vs reality

Historically, most decision making experiments have been done in fully controlled (lab) settings. This makes it possible to isolate interesting dependent variables and control for interfering factors. It does however have one major drawback which is that everyday decision making is not done in a lab. It is not always possible to draw reliable conclusions about human decision making from lab results.

Relevant to the present study is the comparison between computer screen lab experiments and supermarket setting experiments made by Gidlöf et al. (2014). They did an eye-tracking experiment in both the lab and in the supermarket to see if the central bias effect can be found in a natural setting as well as in the lab. The central bias refers to the fact that when looking at items on a computer screen, participants tend to focus more on the items around the center of the

screen. When done in a supermarket setting with items on a shelf, this effect disappeared. The experiment presented in this paper also took place in a supermarket using mobile eye-trackers and can hopefully be a useful addition to the literature on decision making in natural environments. There has been a lot of research done on site in supermarkets (Nordfält, 2007 has a good overview) but eye-tracking is a relatively new addition to that field.

For their article *Shelf Management and Space Elasticity,* Xavier Dreze and others (1995) investigated some in-store strategies to increase sales in a supermarket. Among the things they tried was to reorganize shelves to facilitate what they called "ease-of-shopping". In one of their attempts to create such user-friendly shelves for the study, they organized cereals into types (family, kids, adults) and into subtypes (raisin, bran, fruit, high fiber etc.). They considered this to be the more logical way to organize the shelf, but there was no discussion about why this is more logical. A possible hidden premise in their reasoning here is that consumers know what type, and even subtype, of product they're looking for, e.g. bran flakes, and that putting all bran flakes of different brands together would make it easier to compare and choose the best one.

Organizing the shelf according to subcategories entails separating products of the same brand into different parts of the shelf in order to block them together with similar products (see fig. 2, p. 5). For instance, it is not uncommon to find all ecological options from several different manufacturers blocked together. This inevitably entails a spreading out of the different brands.

Inversely, when organizing a shelf into blocks consisting of all the products from the same brand, that entails a spreading out of products from a certain category (fig. 2, p. 5). On a brand-organized shelf you would have to search every brand-block for their ecological options. For a consumer who prefers ecological products and wants to compare the available options, this requires more effort since in order to compare the ecological products they would have to keep prices and attributes in their working memory for much longer while searching.

Dreze et al. (1995) found that their "user-friendly", category organized, setup resulted in decreased sales for that whole section of the store. They also found that inversely, making some comparisons harder by separating different size packages of the same brand increased sales. They speculate that organizing the shelf according to category made it "too easy" for the consumers to find what they were looking for and thus reduced browsing and impulse purchases. This explanation is echoed in an article by Clement, Astrup & Forsberg (2015) who, in their eye-tracking study, found that placing comparable products from a private label and a strong national brand together decreased total visual attention (fixation time)—presumably because it made the choice easier and decision time shorter. It's an intriguing idea that highlights the possible tension between what's good for the customer, namely being able to easily find what one wants; and what's good for sales, namely keeping the customer browsing the shelves.

When Dreze et al. speculate that the reason for the decrease in sales following that re-organization of the cereal section into categories is that the new layout makes it easier for consumers to swiftly find what they are looking for, they are making a generalization that can be tested. Namely that most if not all consumers approach the shelf with an idea of what category, but not necessarily brand, they are looking for. It could of course be the case that many care more about a certain brand and want to browse its products as in the examples above. Dreze et al. makes no case one way or the other. Making it easy for the first group will not necessarily make it easy for the second, in fact, it seems that they're on the opposite end of the scale.

A more charitable interpretation of Dreze's explanation is that consumers who are looking for a certain category of product will have an easier time finding it if the shelf is organized by category rather than brand. We can also add the corollary, that consumers in the market for products from a certain brand will have a harder time when the products are not organized by brand. This boils down to a simple variable, namely how well the shelf matches a consumer's preferences.

One way to test the explanation suggested by Dreze et al. could be to sort consumers into groups based on how well the shelf layout matches their preferences and then make measurements of cognitive effort. Using eye-tracking, we are in a unique position to actually investigate where Dreze and others have merely speculated. A person's gaze is an arrow pointing not only to what is being looked at but also to where that person's *attention* is directed. It is a subtle but important distinction because it enables us to say something about the decision making process.

Mobile eye-tracking technology is now sophisticated enough to be used to get solid, reliable results outside the laboratory (Holmqvist et al., 2011; Gidlöf, 2014). What is measured is the fixations of a person's gaze on a visual scene. As a person's gaze jumps from point to point on a visual scene, it is important to know when they are actually taking in information. The gaze's fixation needs to have a long enough duration for comprehension to take place. In the case of reading for instance, a word fixation needs to be around 120ms for the reader to comprehend the word (Gidlöf et al., 2013). Such a fixation is called a *dwell.*

Of course, in a supermarket, most of the time one does not read single words but rather looks at products. Just like a word, a product can be a "chunk" of meaningful information. In eye-tracking terms it is called an *Area Of Interest* or *AOI* (see Holmqvist et al., 2011 for details). Since each product often has several facings (i.e. visible packages of the same product) and since for our purposes it does not matter which package is looked at, we have defined an AOI as consisting of all visible packages of the same product. Every time a consumer looks at a product on a shelf, that is one dwell on one AOI. Whenever they return to an AOI they already looked at, that is part of a subset of dwells called *re-dwells.*

In our study, we fitted shoppers in two supermarkets with eye-tracking glasses. Key differences in shelf-layout between the supermarkets made it possible to investigate the relation between visual information organization and decision making and cognitive effort. We also collected preference data covering both brand preference (e.g. leading brands such as Barilla and Yoggi or the ICA store brand) and preferences for several subcategories, such as "light yoghurt", "short pasta" or eco-friendly products. Using this data, we were able divide the participants into a "category-focused group" and a "brand-focused group" where the former tended to value categories significantly higher than brands and vice versa for the latter. We also created a "neutral" group of participants who's preferences did not line up strongly in favor of one or the other. We could then compare the groups, and analyze the eye-tracking and survey data to see what

makes shopping easier and for whom.

We hypothesized that the group of participants who's preferences pertained to certain categories such as ecological alternatives would have an easier time choosing a product when the shelf was organized by that category. Vice versa, we hypothesized that shoppers whose preferences pertained more to certain brands would have an easier time shopping if the shelf was organized by brand. We call this positive relation between preferences and visual information organization (i.e. shelf layout) *congruence.* A preference mismatch such as a category-focused customer at a brand-organized shelf on the other hand is a case of *incongruence.*

In the two supermarkets we tested, there were relevant differences in how the pasta and yoghurt shelves were organized. In supermarket 1, the pasta was organized by category to a much larger extent. For instance, all the ecological products were blocked together in a square. In supermarket 2 the pasta was organized by brand rather than category and the ecological products, to use the same example, were spread out. In the yoghurt section, supermarket 1 had the health options (e.g. "light" or low fat (<1%)) blocked together whereas supermarket 2 had these grouped with other options from the same brand. Depending on the shelf layout and on what the customer is looking for their gaze will have to move in different ways across the scene in order to take in the information necessary to make the decision. Given that a person's gaze is a strong indicator of what visual information they are collecting and considering, eye-tracking is perfectly suited to our experiment.

Overview of the study

Using the eye-tracking data and the self-reported ease-of-shopping data in combination with the self-reported preferences of the participants, we were able to test to what extent a shelf organization that matched a participant's preference made it easier—that is, less cognitively taxing—for them to find what they were looking for.

To do this we conducted the experiment in two supermarket stores with different shelf layouts. We recruited participants outside each supermarket and the participants from the first supermarket also participated at supermarket 2. This enabled us to do between- and within-group comparisons between the two supermarkets; and between-group comparisons at both supermarkets.

The participants were asked to buy food items from three categories: pasta, yoghurt and cereal. Their eye-movements were recorded for the duration of the shopping and upon completion they filled in a survey with questions pertaining to their product preferences and normal shopping routines. The group from the first supermarket was asked to return to the second part of the experiment, one month later, at another supermarket across town. There they—as well as the new recruits at that store—followed the same procedure as in the first store.

Measurements

M1. Self-report based on the survey questions "How easy was it to choose the pasta product?" and "How easy was it to choose the yoghurt product?" on a scale ranging from -5 ("hard") to 5 ("easy").

M2. Time spent looking at the shelf: the amount of time from the first dwell until the last, only counting time spent looking at products on the shelf.

M3. Time spent evaluating the options: a subset of M2 consisting of all dwells from the first to the last dwell on the to-be-chosen product (using the NDSM model: see fig. 1, p. 2).

M4. Total amount of re-dwells, that is every time a participant looked at a product that they had already looked at, from the first dwell to the last.

M5. Amount of re-dwells during the evaluation stage (see M3.)

M6. Number of products looked at (each AOI counted only once).

M7. Decision quality (normalized option quality for the chosen product). Based on product attributes and the self-reported preferences pertaining to those attributes, *option quality* is a normalized value for each product indicating how well each product fits a consumer's preferences. *Decision quality* (DQ) is the ratio of the chosen product's option quality *(Qchosen)* to the highest possible option quality for that consumer (*Qbest)* (the best product given that consumer's preferences). Thus, DQ = Qchosen / Qbest.

Where applicable, values were normalized to take into account the differences in number of products and product facings in the two supermarkets, which can influence how easy it is to find a certain product merely by having more or less products to sift through.

Hypotheses

H1.1
Participants with preferences congruent with the shelf organization will have an easier time choosing their product (measurement M1) than participants with preferences that are incongruent with the shelf organization at the *same supermarket* (between-group comparison).

H1.2
Participants with preferences congruent with the shelf organization at the second supermarket will have an easier time choosing their product (M1) there than the *same participants* in the first supermarket (within-group comparison).

* * *

H2.1
Participants with preferences congruent with the shelf organization will spend less time looking at the products on the shelf (M2) than participants with preferences that are incongruent with the shelf organization at the *same supermarket* (between-group comparison).

H2.2
Participants with preferences congruent with the shelf organization at the second supermarket will spend less

time looking at the products on the shelf (M2) there than the *same participants* in the first supermarket (within-group comparison).

* * *

H3.1
Participants with preferences congruent with the shelf organization will have a shorter evaluation stage (M3) than participants with preferences that are incongruent with the shelf organization at the *same supermarket* (between-group comparison).

H3.2
Participants with preferences congruent with the shelf organization at the second supermarket will have a shorter evaluation stage (M3) there than the *same participants* in the first supermarket (within-group comparison).

* * *

H4.1
Participants with preferences congruent with the shelf organization will have fewer re-dwells (M4) than participants with preferences that are incongruent with the shelf organization at the *same supermarket* (between-group comparison).

H4.2
Participants with preferences congruent with the shelf organization at the second supermarket will have fewer re-dwells (M4) there than the *same participants* in the first supermarket (within-group comparison).

* * *

H5.1
Participants with preferences congruent with the shelf organization will have fewer re-dwells during the evaluation stage (M5) than participants with preferences that are incongruent with the shelf organization at the *same supermarket* (between-group comparison).

H5.2
Participants with preferences congruent with the shelf organization at the second supermarket will have fewer re-dwells during the evaluation stage (M5) there than the *same participants* in the first supermarket (within-group comparison).

* * *

H6.1
Participants with preferences congruent with the shelf organization will focus on fewer products (M6) than participants with preferences that are incongruent with the shelf organization at the *same supermarket* (between-group comparison).

H6.2
Participants with preferences congruent with the shelf organization at the second supermarket will focus on fewer products (M6) there than the *same participants* in the first supermarket (within-group comparison).

* * *

H7.1
Participants with preferences congruent with the shelf organization will make better purchase decisions (higher decision quality, see M7) than participants with preferences that are incongruent with the shelf organization at the *same supermarket* (between-group comparison).

H7.2
Participants with preferences congruent with the shelf organization at the second supermarket will make better purchase decisions (M7) there than the *same participants* in the first supermarket (within-group comparison).

These hypotheses leave out two possible comparisons: The reversed within-group comparison across supermarkets; and the between-group comparison across supermarkets. The first is left out because there are confounding factors and effects that work in the same direction as the investigated effect. The second is left out because neither of the two main variables would be kept constant.

2 Materials and methods

The study took place in two supermarkets in Lund, Sweden. The supermarkets were situated in the east and west part of town respectively:

ICA Kvantum Malmborgs Tuna (supermarket 1)
ICA Kvantum Malmborgs Mobilia (supermarket 2)

They both carried more or less the same product range but supermarket 2 was bigger and with bigger shelves and more more facings (336 vs 211 facings for pasta; 202 vs 161 for yoghurt). Importantly the shelves were also organized differently with supermarket 1 having shelves organized mainly by category and sub-category (e.g. ecological products; healthier options etc.) while supermarket 2 had shelves organized mainly by brand (fig. 2).

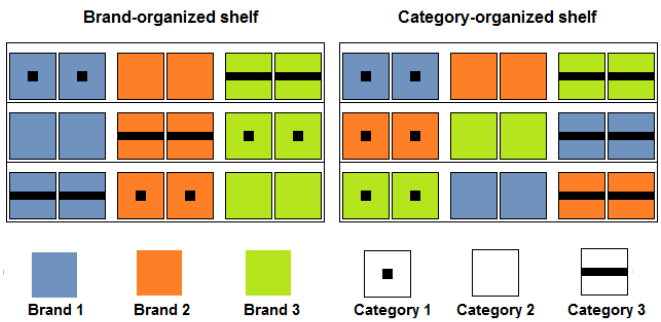

Fig. 2. A simplistic visualization of the difference between a shelf organized by brand and a shelf organized by category.

Participants

50 participants (mean age: 21.4, *SD:* 2.6; 18 female) were recruited at supermarket 1. 38 of them also joined the second part of the experiment at the second supermarket (mean age: 21.5, *SD: 2.8, 16 female).* There another 26 participants were recruited (mean age: 22.8, *SD:* 5.1, 13 female). All participants had normal or corrected to normal eyesight.

In the first part of the study, participants received a SEK100 gift check to cover the costs of pasta, yoghurt and cereal purchases based on the sum of the most expensive products in each category. In the second part of the study, participants again received a SEK100 gift check and the ones who had also partaken in the first part were rewarded with a cinema ticket voucher for making the extra trip.

Stimuli & design

Three product categories available in different sections in the supermarkets were chosen as prompted purchases; pasta, yoghurt and cereal. In the case of each category, the shelf display was divided into AOIs each of which consisted of the package facings of all items of each product of each brand (e.g. AOI 1: ICA light strawberry yoghurt, AOI 2: ICA blueberry yoghurt, AOI 3: Yoggi blueberry yoghurt etc.).

Apparatus

Each participant was fitted with SMI eye tracking glasses which record (binocularly) at 30hz. The SMI glasses were connected to—and their data recorded on—a laptop placed in a backpack worn by the participant during their shopping session. Three pairs of glasses were used.

Before each recording, the eye tracker was calibrated using a 1-point calibration setup. In some cases a 3-point calibration was needed. At the end of each recording, the participants were asked to focus on and follow the experiment leader's fingertip with their gaze while keeping their head still. This was done to ensure reliable recording throughout each session (to be retrospectively confirmed during data analysis). SMI's Iview software was used for the recording.

Procedure

Participants were asked to buy one pasta item, one yoghurt item and one cereal item in addition to their intended, regular shopping. There was no time constraint and no further instructions. Once they were done with their shopping trip, their pasta, yoghurt and cereal choices were noted and they were asked to fill in two questionnaires; one having to do with their shopping routines and one having to do with product preferences (e.g. brand, sugar content etc.).

Questionnaires

The questionnaires were designed with several different post-experiment analyses in mind so not everything on them was of relevance to this paper.

The familiarity questionnaire consisted of questions designed to assess the participant's familiarity with the supermarkets and their products. Of main importance to the present study, it also enquired about how easy the participant found choosing their products (one answer per category ranging from -5 (hard) to +5 (easy)).

The preference questionnaire listed attributes such as price, country of origin, whether it was ecological etc. It also listed every brand available in the different categories. The participants were asked how each attribute affected them in their choice of product ranging from -5 (negatively) to +5 (positively). This data was used for the present analysis to group participants according to their preferences towards categories and brands. It was also used to calculate *option quality* for each product and ultimately *decision quality* (for the chosen product). The participants also signed a standard experiment consent form before the experiment.

Analysis

Only the pasta and yoghurt data was analyzed for the present study. For the analysis of the pasta data, a total of 20 participants had to be excluded due to either loss of eye-tracking data or for buying pasta from other parts of the store than the assigned shelf. For the yoghurt data analysis, 14 participants had to be excluded for the same reasons.

The participants were divided into preference groups based on their responses to the preference questionnaire. In the analysis of the pasta data a participant was considered category-focused (the category being ecological products in this case) based on their average ratings of the following product attributes: *"Ekologisk", "Lantmännen-märkt", "EU-märkt ekologisk vara", "Klimatdeklarerad" and "Kravmärkt"* (each of them Swedish and/or European standards for ecological products, see KRAV-märket och andra miljömärkningar (2015) for details).

In the analysis of the yoghurt data a participant was considered category-focused (the category being healthier yoghurt options in this case) based on their average ratings of the following product attributes: *"Nyckelhålet-märkt"* (The Keyhole, 2015), *"Light", "Low sugar", Low fat", "Low calorie"*.

To be considered category-focused the following conditions must be met (possible ratings for any attribute ranges from -5.0 to +5.0.):

(i) The average rating for the category attributes above must exceed 3.5.

(ii) The maximum rating for brand attributes may not exceed 3.0.

To be considered brand-focused the following conditions must be met:

(i) At least one brand must be rated 4.0 or higher.

(ii) The average rating for the category attributes above must not exceed 3.0.

Any participants who did not meet any of the above criteria were put in the *control group*. Their preferences were not considered congruent with either a brand-organized or a category-organized shelf.

The division admits of no overlap and so each participant is in one and only one of the three groups. Since the division is done separately for pasta and yoghurt respectively however, it is possible for the same participant to be part of one group for the pasta data analysis and another for the yoghurt data analysis. For a discussion about the thresholds used, see the *Definitions* section in *Discussion*.

The raw eye-tracking data was encoded using semantic gaze mapping in the SMI BeGaze software. This was done by manually connecting the eye's gaze in each frame at the relevant supermarket shelf to a high resolution image of that shelf. The image was then mapped with AOIs covering each product's facings. The data was then converted from individual focuses in each frame into dwells. A dwell was defined as a sequence of fixations within the same AOI that lasted for more than 120ms. This threshold was decided upon using the definition of a dwell from Gidlöf et al. (2013). Finally the data for each participant was divided into an orientation stage and an evaluation stage where the dividing line was the first dwell on the AOI representing the product that was ultimately chosen.

Option quality for each product was calculated using a list of each product's attributes and the data from the preference questionnaire of the same attributes. For instance, for someone looking for low fat yoghurt, the yoghurt with

the lowest fat content would have the highest option quality score for that attribute. Comparing all attributes in this way gives each product a total option quality rating for each participant. The option quality for the chosen product divided by the maximum possible option quality for that choice is called the *decision quality* and is a normalized rating of how well the shopper managed to match their stated preferences in their purchase decision.

3 Results

Preference congruence groups

The division of participants into groups according to their preferences resulted in the distribution found in table 1 below.

Table 1. Number of participants in each preference group with the supermarket 2 total including both returning participants from supermarket 1 and those recruited at supermarket 2 (last column).

Pasta	**Supermarket 1**	**Supermarket 2 (total)**	**Recruited at Supermarket 2**
Category	**10**$^{+}$	13^{-}	6
Brand	20^{-}	**28**$^{+}$	12
Neutral	20	22	7

Yoghurt	**Supermarket 1**	**Supermarket 2 (total)**	**Recruited at Supermarket 2**
Category	**14**$^{+}$	15^{-}	6
Brand	24^{-}	**30**$^{+}$	10
Neutral	12	18	9

In table 1, groups marked with a plus sign are those who's preferences (table row) are in alignment with the shelf organization of that kind of product in that particular supermarket (table column). In the section to follow, the word *congruent* will be used to denote those groups and each table will contain a column (*"Con."*) with either a plus or a minus sign. Groups marked with a minus sign are the *incongruent* groups who's preferences conflict with the organization of that store's shelves. Finally, unmarked groups are *neutral* since their preferences neither align nor conflict with that store's shelf organization. Since different categories were used for pasta and yoghurt it was possible for a participant to be in different groups for the two product types (e.g. category-focused for pasta but brand-focused for yoghurt). 10 of the participants were in the category-focused groups for both pasta and yoghurt; 18 participants were in both brand-focused groups and 12 were in the control-groups for both pasta and yoghurt.

All paired, within-group, t-tests reported only take into account those participants of a certain group that visited *both* supermarkets. For instance, of the 20 brand-focused participants recruited at the first supermarket 16 also took part in the experiment at the second supermarket. A paired within-group t-test of that group would then only contain those 16 participants' data from both supermarkets.

Ease-of-shopping

Ease-of-shopping—a term borrowed from the aforementioned article by Dreze et al. (1995)—is here defined as a participant's rating, from -5 to 5, of how easy it was to choose a particular type of product (e.g. pasta, or yoghurt).

Bootstrapped calculated mean averages (Kaplan, 2015) indicated significant differences between congruent groups on the one hand and both neutral and incongruent groups on the other. This held for pasta in both supermarket 1 (fig. 3) and 2 and for yoghurt in supermarket 2 (table 2). Between-group t-tests on the data for pasta supported the tendencies in the bootstrapped data ($t(28) = 2.20$, $p = 0.04$ in the first supermarket and $t(31) = -1.77$, $p = 0.087$ in the second). A paired, within-group, t-test of the brand-focused pasta group from supermarket 1 (incongruent at supermarket 1, congruent at supermarket 2) did not point to any difference. For the brand-focused yoghurt group from supermarket 1 (incongruent at supermarket 1, congruent at supermarket 2) however, a paired within-group t-test did indicate an effect ($t(19) = -2.08$, $p = 0.05$) (fig. 6). A paired within-group t-test of the control group from supermarket 1 indicated that they had a slightly more difficult time at supermarket 2, but the results are not reliable (fig. 7). A linear fitted model for the alignment of preferences did not indicate any correlations between preference alignment and ease-of-purchase for pasta (fig. 4) and only a weak correlation for yoghurt ($r(48) = 0.32$, $p = 0.022$; fig. 5). All correlations reported are Pearson's *r*.

Table 2. Bootstrapped mean average values for self-reported ease-of-shopping for pasta and yoghurt with a 90% confidence interval.

Pasta *Supermarket 1*	**Con.**	**Mean**	**[5% – 95%]**
Category-focused	+	3.40	2.54 – 4.26
Brand-focused	-	1.70	0.36 – 3.05
Control-group		1.95	0.87 – 3.03
Supermarket 2			
Category-focused	-	2.47	1.27 – 3.67
Brand-focused	+	3.03	2.28 – 3.79
Control-group		2.41	1.51 – 3.31
Yoghurt *Supermarket 1*			
Category-focused	+	1.00	-0.38 – 2.38
Brand-focused	-	1.88	0.85 – 2.90
Control-group		1.99	-0.22 – 4.21
Supermarket 2			
Category-focused	-	2.13	0.82 – 3.43
Brand-focused	+	2.90	2.19 – 3.60
Control-group		2.40	1.15 – 3.64

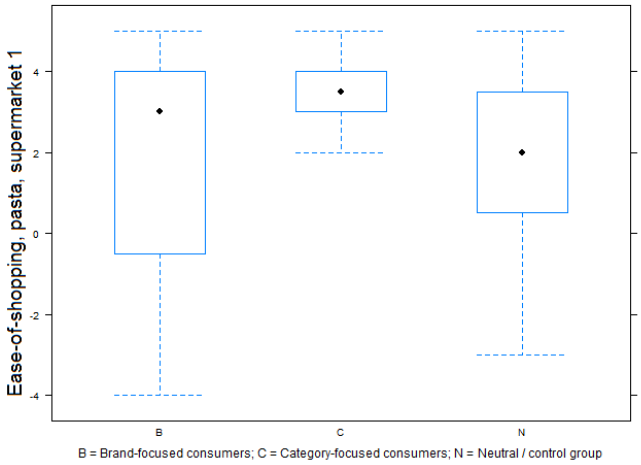

Fig. 3. Boxplot showing self-reported ease with which the participants chose their pasta product in supermarket 1 on a scale ranging from -5 (very difficult) to +5 (very easy).

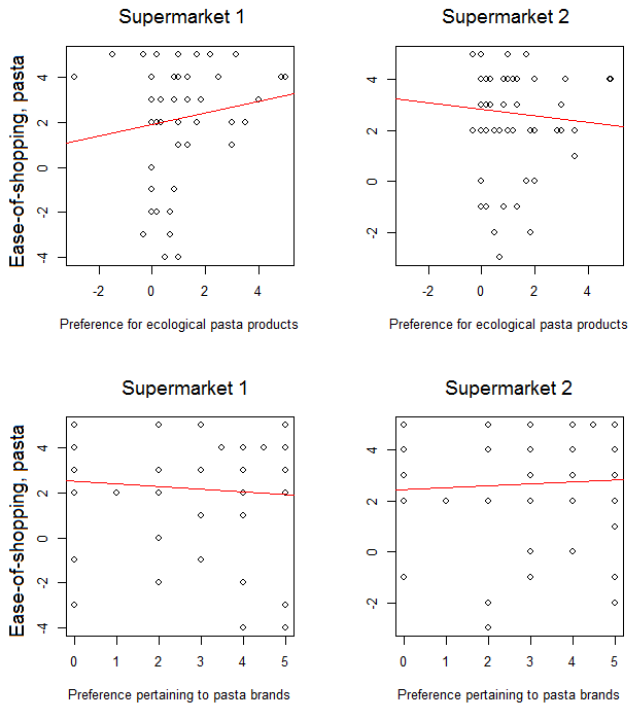

Fig.4. Scatterplots of self-reported ease-of-shopping (pasta shelf) for each supermarket and preference type combination. The red lines represent linear fitted models of the data that show the correlation coefficient (Pearson's *r*). The top left and bottom right graphs show the congruent conditions; the top right and bottom left graphs show the incongruent conditions.

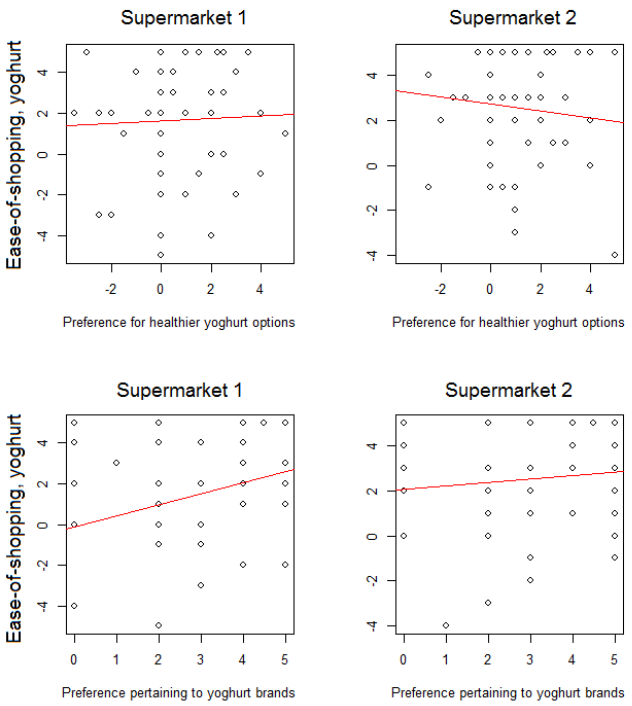

Fig. 5. Scatterplots of self-reported ease-of-shopping (yoghurt shelf) for each supermarket and preference type combination. The red lines represent linear fitted models of the data that show the correlation coefficient (Pearson's *r*). The top left and bottom right graphs show the congruent conditions; the top right and bottom left graphs show the incongruent conditions.

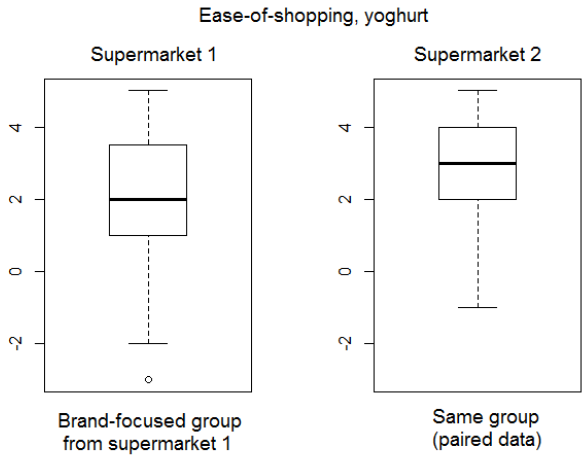

Fig. 6. Boxplots showing a paired within-group comparison of self-reported ease-of-shopping (choosing a yoghurt product) for the brand-focused participants from supermarket 1 that also participated at supermarket 2.

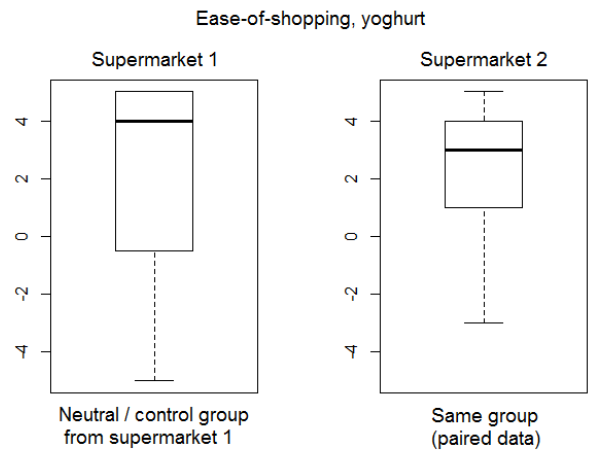

Fig. 7. Boxplots showing a paired within-group comparison of self-reported ease-of-shopping (choosing a yoghurt product) for the neutral control group of participants from supermarket 1 that also participated at supermarket 2.

Dwell time on the shelf

The total dwell time was measured from the first dwell on the product shelf to the last dwell on the shelf.

Bootstrapped mean average calculations suggest that preference-congruent groups are faster in both supermarkets for pasta but not for yoghurt in supermarket 1 (table 3; fig. 8 & 9). A between-group t-test (independent 2-group) comparing congruent and incongruent groups' total dwell time at the yoghurt shelf at supermarket 2 indicated some significance ($t(12) = -1.93$, $p = 0.077$).

A within-group (paired) t-test on the supermarket 1 control (neutral) group's total dwell time for pasta at supermarket 1 and 2 respectively indicated that they were faster at supermarket 1 ($t(8) = -2.57$, $p = 0.03$) (fig. 10).

Table 3. Bootstrapped mean average values for total dwell time (ms) for pasta and yoghurt with a 90% confidence interval.

Pasta *Supermarket 1*	**Con.**	**Mean**	**[5% – 95%]**
Category-focused	+	25331	7081 – 43579
Brand-focused	-	33539	16815 – 50262
Control-group		25380	14083 – 36676
Supermarket 2			
Category-focused	-	50675	15352 – 85996
Brand-focused	+	46399	31550 – 61248
Control-group		38035	23909 – 52161
Yoghurt *Supermarket 1*			
Category-focused	+	43672	17214 – 70130
Brand-focused	-	26846	15897 – 37795
Control-group		22936	4657 – 41215
Supermarket 2			
Category-focused	-	56703	28624 – 84783
Brand-focused	+	29394	22645 – 36143
Control-group		32167	18319 – 46431

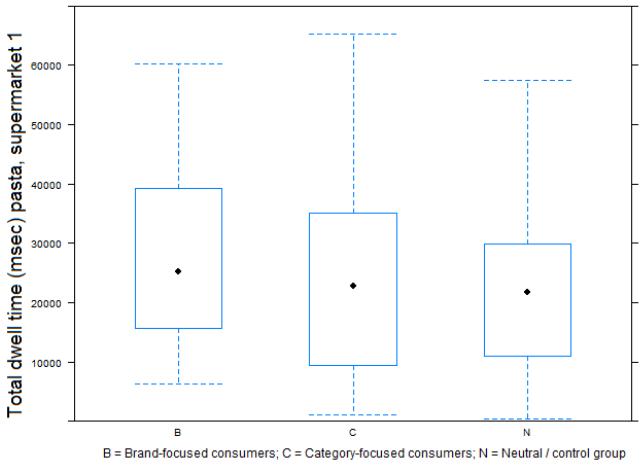

Fig. 8. Boxplot showing total dwell time in milliseconds in supermarket 1 (pasta shelf) for the three preference groups. In this supermarket the category-focused (middle) group is the congruent group.

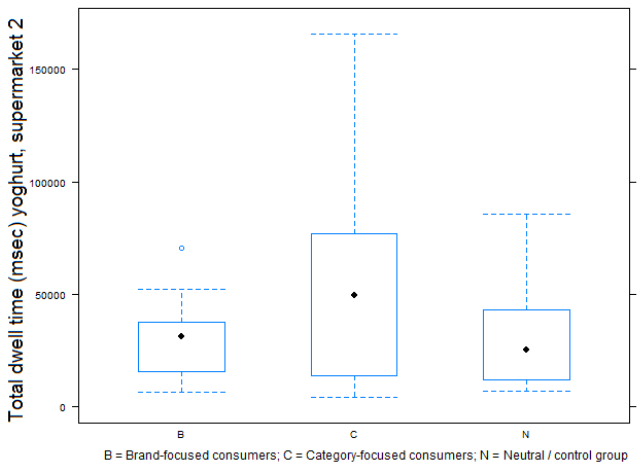

Fig. 9. Boxplot showing total dwell time in milliseconds in supermarket 2 (yoghurt shelf) for the three preference groups. In this supermarket the brand-focused (left) group is the congruent group.

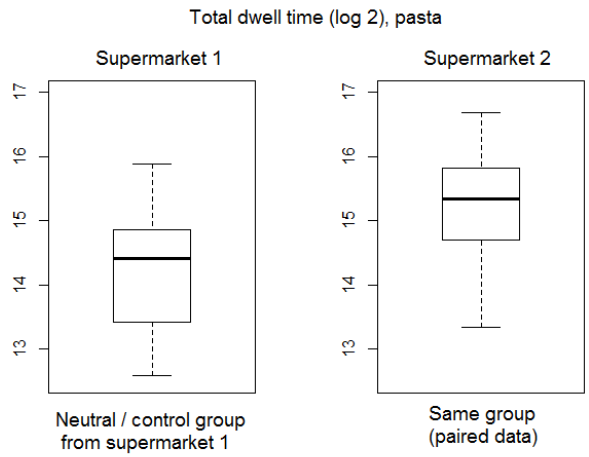

Fig. 10. Boxplots showing a paired within-group comparison of total dwell time (on a log 2 scale) for pasta for the neutral control group of participants from supermarket 1 that also participated at supermarket 2.

The evaluation stage is defined as the interval between the first and the last dwell on the product that was ultimately chosen.

Bootstrapped mean averages showed no significant between-group differences in dwell time during the evaluation stage for the groups at supermarket 1 (table 4). In supermarket 2, the incongruent group tended to be slower than the neutral group for pasta (table 4, fig. 11). At the yoghurt shelf at supermarket 2, the congruent group tended to be faster than the incongruent group.

Table 4. Bootstrapped mean average values for dwell time (ms) during the evaluation stage for pasta and yoghurt with a 90% confidence interval.

Pasta *Supermarket 1*	**Con.**	**Mean**	**[5% – 95%]**
Category-focused	+	20273	2997 – 37548
Brand-focused	-	20805	4220 – 37389
Control-group		19295	6994 – 31595
Supermarket 2			
Category-focused	-	46967	9153 – 84780
Brand-focused	+	33879	17571 – 50187
Control-group		19277	11162 – 27389
Yoghurt *Supermarket 1*			
Category-focused	+	13635	5154 – 22117
Brand-focused	-	17245	6600 – 27890
Control-group		13605	-2316 – 29525
Supermarket 2			
Category-focused	-	39195	14133 – 64256
Brand-focused	+	18651	11986 – 25315
Control-group		17988	6534 – 29443

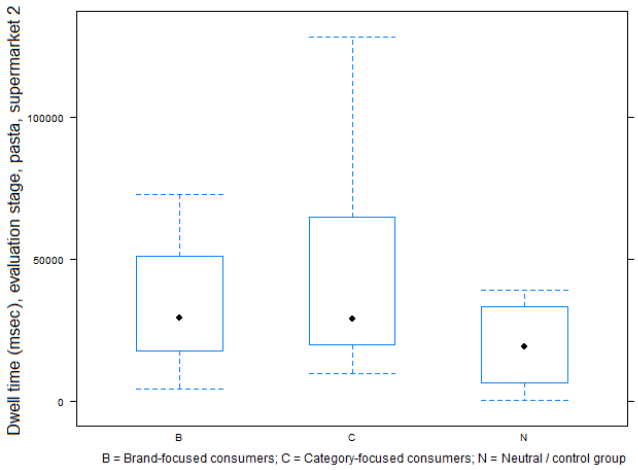

Fig. 11. Boxplot showing evaluation dwell time in milliseconds in supermarket 2 (pasta shelf) for the three preference groups. In this supermarket the brand-focused (left) group is the congruent group.

Re-dwells

Re-dwells are defined as returns to AOIs (e.g. in the following AOI sequence **A1**, A2, A3, **A1**, only the last A1 dwell is a re-dwell) already looked at.

A bootstrap calculation of the mean average number of re-dwells indicated that the congruent group at supermarket 1 made more re-dwells than the incongruent group for pasta (table 5).

At the yoghurt shelf in supermarket 2, the incongruent group tended to make more re-dwells than both the congruent and the neutral (control) group (fig. 12). In the evaluation stage re-dwell data the only somewhat significant result was the indication that the incongruent group made more re-dwells than the neutral group ($t(17) = 1.82$, $p = 0.087$) (fig. 13).

Table 5. Bootstrapped mean average values for total number of re-dwells for pasta and yoghurt with a 90% confidence interval.

Pasta *Supermarket 1*	**Con.**	**Mean**	**[5% – 95%]**
Category-focused	+	31.4	2.5 – 60.1
Brand-focused	-	20.7	2.2 – 39.3
Control-group		31.1	3.0 – 59.1
Supermarket 2			
Category-focused	-	53.8	-2.3 – 110.0
Brand-focused	+	24.9	10.9 – 39.0
Control-group		17.9	10.0 – 24.7
Yoghurt *Supermarket 1*			
Category-focused	+	14.3	5.0 – 23.7
Brand-focused	-	17.2	5.5 – 28.8
Control-group		14.2	-5.9 – 34.3
Supermarket 2			
Category-focused	-	28.3	11.1 – 45.6
Brand-focused	+	16.7	9.9 – 23.4
Control-group		15.7	4.3 – 27.1

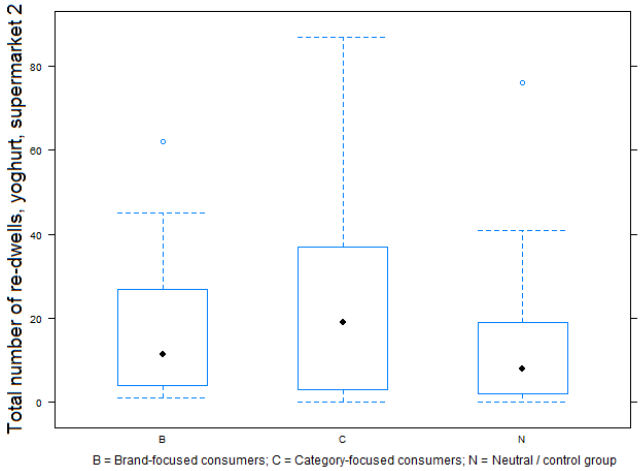

Fig. 12. Boxplot showing total number of re-dwells in supermarket 2 (yoghurt shelf) for the three preference groups. In this supermarket the brand-focused (left) group is the congruent group.

Table 6. Bootstrapped mean average values for number of re-dwells during the evaluation stage for pasta and yoghurt with a 90% confidence interval.

Pasta *Supermarket 1*	**Con.**	**Mean**	**[5% – 95%]**
Category-focused	+	18.0	4.6 – 31.5
Brand-focused	-	15.4	4.8 – 26.1
Control-group		18.1	7.5 – 28.6
Supermarket 2			
Category-focused	-	26.3	9.5 – 43.0
Brand-focused	+	18.2	8.2 – 28.2
Control-group		11.6	6.6 – 16.6
Yoghurt *Supermarket 1*			
Category-focused	+	4.2	0.8 – 7.6
Brand-focused	-	6.0	3.2 – 8.7
Control-group		3.3	2.1 – 4.5
Supermarket 2			
Category-focused	-	7.3	4.4 – 10.1
Brand-focused	+	5.8	3.8 – 7.9
Control-group		5.4	3.1 – 7.8

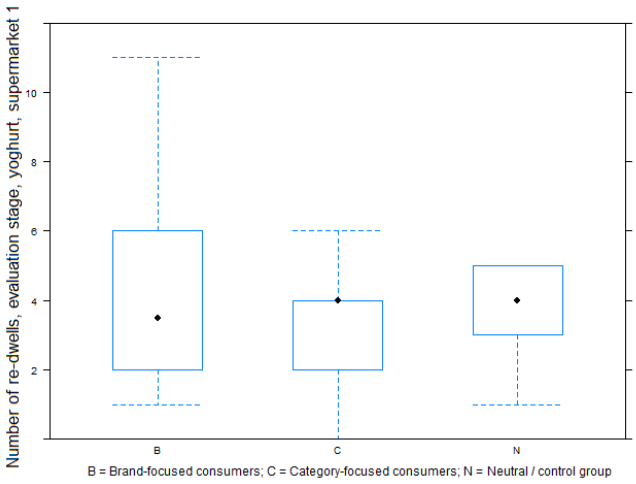

Fig. 13. Boxplot showing the number of re-dwells during the evaluation stage in supermarket 1 (yoghurt shelf) for the three preference groups. In this supermarket the category-focused (middle) group is the congruent group.

Number of products focused on

The number of different products focused on / looked at, from the first AOI dwell on the shelf to the last where each AOI represents all facings of one product.
Comparisons of the bootstrapped mean averages of the different groups suggested some small effects on the number of products looked at (table 7). A fitted linear model of the effect suggested only slight and statistically unreliable correlations (fig. 14) between stronger congruent preferences and a decrease in number of products looked at.

Table 7. Bootstrapped mean average values for number of products focused on for pasta and yoghurt with a 90% confidence interval.

Pasta *Supermarket 1*	**Con.**	**Mean**	**[5% – 95%]**
Category-focused	+	25.3	15.5 – 35.2
Brand-focused	-	31.7	24.1 – 39.3
Control-group		28.6	21.8 – 35.4
Supermarket 2			
Category-focused	-	34.9	22.3 – 47.6
Brand-focused	+	31.4	22.3 – 40.4
Control-group		27.6	21.4 – 33.8
Yoghurt *Supermarket 1*			
Category-focused	+	33.3	18.1 – 48.5
Brand-focused	-	23.3	15.7 – 31.0
Control-group		23.9	11.0 – 36.8
Supermarket 2			
Category-focused	-	35.4	23.1 – 47.7
Brand-focused	+	26.1	21.4 – 30.1
Control-group		27.3	18.8 – 32.8

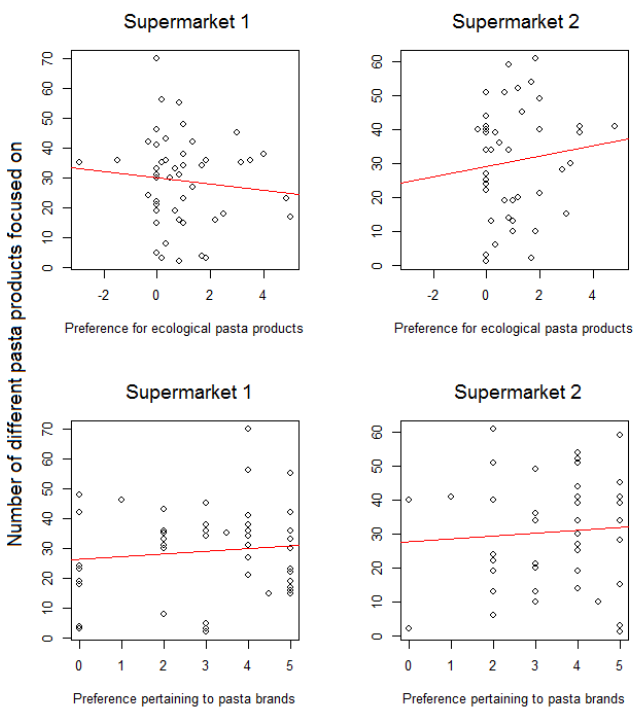

Fig. 14. Scatterplots of number of different pasta products focused on for each supermarket and preference type combination. The red lines represent linear fitted models of the data that show the correlation coefficient (Pearson's *r*). The top left and bottom right graphs show the congruent conditions; the top right and bottom left graphs show the incongruent conditions.

Decision quality

The *decision quality* is based on the *option quality* value for the chosen product and divided by the maximum possible option quality for that choice for each participant (see the *Analysis* section in *Materials and Methods*). Bootstrapped mean averages of decision quality indicated that the congruent group at supermarket 1 made slightly worse pasta purchases given their self-reported preferences than did the incongruent group ($t(16) = -1.60$, $p = 0.12$) (table 8; fig. 16). Between-group t-test showed the same tendency in supermarket 2 where the congruent group made worse decisions than both the incongruent ($t(22) = -1.50$, $p = 0.15$) and the neutral ($t(25) = -1.69$, $p = 0.10$).

A fitted linear model of the effect suggested only a slight, unreliable, inverse correlation (fig. 15) between alignment of preferences and the quality of pasta purchase decisions. That is, congruence between preferences and shelf organization correlates with worse purchase decisions. Using such a fitted model on the yoghurt data (fig. 17) yielded contradictory results with a modest positive correlation between congruence and decision quality in supermarket 1 ($r(38) = 0.32$, $p = 0.04$) but inverse in supermarket 2 ($r(48) = 0.36$, $p = 0.011$).

Table 8. Bootstrapped mean average values for decision quality for pasta and yoghurt with a 90% confidence interval.

Pasta *Supermarket 1*	**Con.**	**Mean**	**[5% – 95%]**
Category-focused	+	0.55	0.27 – 0.83
Brand-focused	-	0.79	0.66 – 0.92
Control-group		0.71	0.46 – 0.96
Supermarket 2			
Category-focused	-	0.65	0.38 – 0.91
Brand-focused	+	0.73	0.60 – 0.86
Control-group		0.64	0.51 – 0.77
Yoghurt *Supermarket 1*			
Category-focused	+	0.56	0.38 – 0.75
Brand-focused	-	0.43	0.34 – 0.53
Control-group		0.49	0.30 – 0.69
Supermarket 2			
Category-focused	-	0.48	0.37 – 0.59
Brand-focused	+	0.38	0.31 – 0.46
Control-group		0.50	0.38 – 0.62

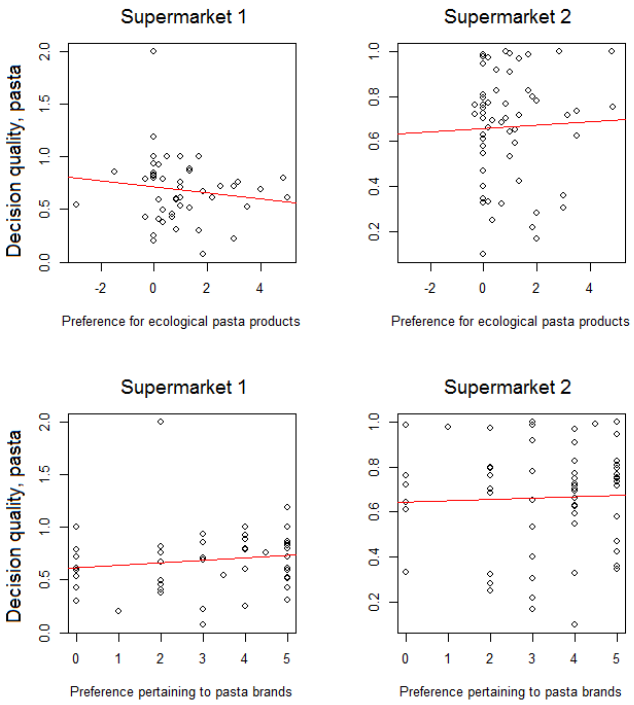

Fig. 15. Scatterplots of decision quality (chosen pasta product) for each supermarket and preference type combination. The red lines represent linear fitted models of the data that show the correlation coefficient (Pearson's *r*). The top left and bottom right graphs show the congruent conditions; the top right and bottom left graphs show the incongruent conditions.

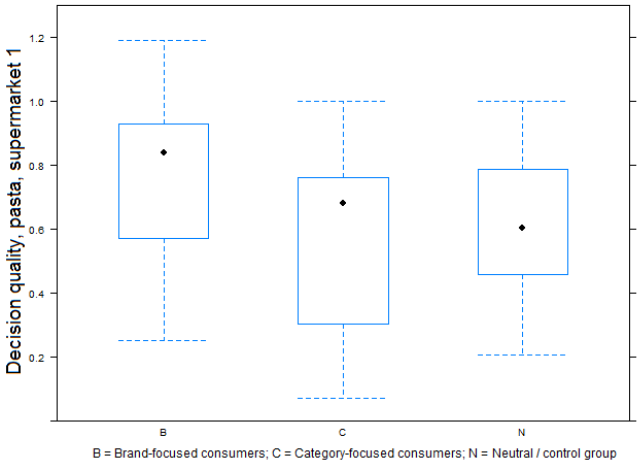

Fig. 16. Boxplot showing decision quality (chosen pasta product) in supermarket 1 for the three preference groups. In this supermarket the category-focused (middle) group is the congruent group.

4 Discussion

There are two lines of inquiry that are of most interest given the hypotheses:

1. Between-group comparisons of participants in the *congruent* and *incongruent* groups in both supermarkets. That is, comparing the data of those for whom the shelf organization should be a support in finding and choosing their product on the one hand with those for whom the organization should interfere with their decision making on the other. Those groups can also be compared with the control group.

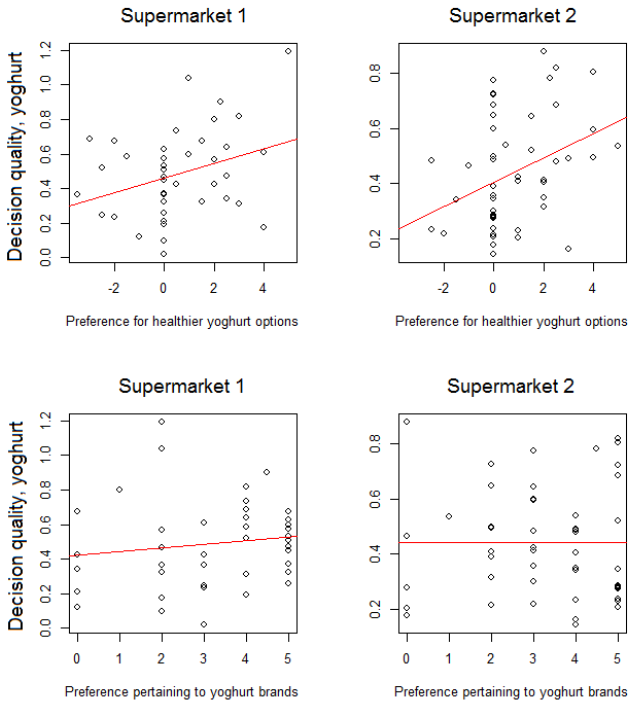

Fig. 17. Scatterplots of decision quality (chosen yoghurt product) for each supermarket and preference type combination. The red lines represent linear fitted models of the data that show the correlation coefficient (Pearson's *r*). The top left and bottom right graphs show the congruent conditions; the top right and bottom left graphs show the incongruent conditions.

2. Within-group comparisons of the returning participants' data from the first and the second supermarket. That is, comparing the data from the two supermarkets where participants' preferences are *congruent* in one and *incongruent* in the other or where they are in the control group and *neutral* in both stores.

A guided tour of shelves and results

At this point it is useful to recap the setting of the experiment and the division into groups before venturing into a discussion of the results. In the first supermarket, the pasta and yoghurt products are organized by category (see fig. 2, p.5). At the first supermarket's pasta shelf this entails blocking of categories such as short (e.g. penne) and long pasta (e.g. linguine) and a block of ecological items from several different brands. This eco category being of special interest for the present study since it presents a useful difference between the layouts of the two supermarkets.

Looking at the data from the pasta shelf in supermarket 1, we can see that the participants in the eco-focused group—who's preferences align well with the organization of the shelf—seemed to have an easier time finding and choosing their product than did the more brand-focused participants.

Looking first at the self-reported ease-of-shopping data, it is clear that the eco-focused participants enjoyed a much easier time choosing their product than did the brand-focused ($t(28 = 2.21$, $p = 0.036$) and control ($t(28) = 2.18$, $p = 0.038$) groups (in accordance with hypothesis H1.1, p. 4). This data is based on a survey they answered immediately after completing their shopping, so it should be a reliable report of their subjective experience. As we saw in the *Results* section it is also in line with the more objective measures discussed below.

At the same shelf, the eco-focused group also spent a lot less time *looking at products* (H2.1, p.4) but only slightly less time evaluating the different options (H3.1, p.5). Their dwell times did not differ very much from those of the control group which could indicate a couple of things: perhaps category organization is, as Drèze et al (1995) suggest, the common-sense organization that most participants are familiar with. It could also be indicative of non-brand organization being a hindrance to those participants who are more brand-focused in their shopping since those participants were slower than both the other groups.

Remaining at the pasta shelf in the first supermarket, the brand-focused group had fewer re-dwells than did both the eco-focused and the control group. This was true of the total as well as the evaluation stage re-dwell data. This is a bit surprising (it contradicts hypothesis H4.1 and H5.1, see p.5) since it could be indicative of them having an *easier* time choosing a product whereas we just saw that on average they themselves reported the opposite. They did however look at *more products* (H6.1) than the eco-focused and the control groups. They were, in other words, looking at a lot of options but not returning to very many of them. Having spent more time and looked at more items, they also ended up making the best decisions ($t(16) = 1.60$, $p = 0.12$) in that they bought products that matched their stated preferences well (H7.1).

As we've seen, fewer re-dwells during the evaluation stage indicates less effort going into the actual choosing of a product (Russo & Leclerc, 1994). Still, according to our data this seems to correlate with good decision making. Perhaps this can be explained with *fast and frugal heuristics* as detailed in the article by Gigerenzer & Goldstein (1996) where they—following Herbert Simon's notion of satisficing—show that simple "Take The Best" algorithms often outperform more complex inference procedures in making fast and accurate decisions. All in all we found an inverse correlation between decision quality and preference congruence for pasta products (see fig. 15 in *Results*).

Moving on to the yoghurt shelf in the first supermarket. It too is a shelf organized by categories. In this case for instance, fruit yoghurt or "natural" yoghurt and which also has a section with healthier options such as "light", "mini" low-fat and low-sugar alternatives. This healthy category is the focus of this part of the study. We looked at the participants' preferences pertaining to the relevant product attributes and again (afterwards, for the analysis) divided them into a health-focused group, a brand-focused group and a neutral control group.

Again, note that although it is the same participants visiting both shelves during the same shopping trip, and that they are divided into category-, brand- and control-groups, the same participant can be in different groups for pasta and yoghurt. An eco-person at the pasta shelf can be in the control group for yoghurt for example.

Looking at the data for the yoghurt shelf in the first supermarket, we can see that the story looks a little different than it did at the pasta shelf. Here the category-focused group (healthier yoghurt options) reported having a harder time (H1.1); spent the most time looking at the shelf (H2.1); less time evaluating the options (H3.1); had fewer re-dwells both total (H4.1) and during evaluation (H5.1); looked at more products (H6.1) and made worse decisions (7.1) than the other groups. In other words, it was pretty much the other way around for this category-group compared to the eco-category pasta-group.

Perhaps the "healthier option" category was not as well defined as eco-pasta. Perhaps its items were not as conspicuous as the mostly green eco-pasta packaging. There are however tendencies that should be noted. For both pasta and yoghurt spending more time looking at more products but less time evaluating (fewer) options seems to go hand in hand with finding the process harder while also making worse purchase decisions. It is an interesting observation and one that retailers should perhaps be aware of if they want to know how to make their shoppers satisfied with their shopping.

We move now to the second supermarket, where the shelves are organized by brand and where items in the eco and healthier categories are spread out accordingly. That is, each brand has its eco and healthier options together with their other products which means that someone who is in the market for, say, an ecological pasta will have to cover a lot more shelf area to see and be able to compare the different eco options.

With the data from the second supermarket, we can do both between- and within-group comparisons since a subset of the participants there were also part of the first part of the study at the first supermarket. For the within-group comparisons, we only used the preference survey results from the *first supermarket* so that they stay within the same preference group for both supermarkets. This decision rests on the assumption that their preferences pertaining to ecological and healthier products as well as different brands did not change too much in the few weeks between the experiments. Apart from making within-group comparisons more straightforward, this also means that all participants (whether recruited at supermarket 1 or 2) had their preferences recorded at their *first* experiment session.

The pasta shelf at the second supermarket is organized by brand. It consists of big product blocks of different colors depending on the brand making up the block (see fig. 2, p.5). This makes it usefully different from the pasta shelf in the first supermarket and it is to this difference we now turn.

Looking first at the ease-of-shopping data, the within-group comparisons show that in accordance with our hypothesis (H1.2), the eco-focused group, who had the easier time in the first supermarket, reported having a harder time at the second supermarket. So did the control group to some extent, but the difference there was insignificant. Worth noting here is the role of familiarity. Since the participants from the first supermarket were all recruited there and since the stores are located quite far apart, we should expect them to be less (if at all) familiar with—and therefore have a harder time finding their way around—the second supermarket. For the brand-focused group the results were the opposite. They reported having an easier time in the second supermarket which also supports our hypothesis (H1.2). Among the three groups in supermarket 2, the brand-focused group reported having an easier time choosing a product than did both the eco-focused and the control groups (H1.1).

When instead of looking at the groups we check for correlations between preferences pertaining to categories and brands respectively on the one hand and ease-of-shopping on the other, the picture becomes clear: congruence between preferences and shelf organization leads to (an experience of) easier shopping at the pasta shelf (see fig. 4 in *Results*).

Looking at the dwell time data, we see familiar tendencies. A within-group comparison of the eco-focused

group's data showed that they were much slower in the second supermarket: both in total time looking at the shelf (H2.2) and evaluation stage time (H3.2). Again, familiarity might well be a factor here. Comparing the groups, we see that the brand-focused group spent less total time (H2.1) and time evaluating (H3.1) options than the eco-group.

The within-group comparisons for number of products focused, total re-dwells and evaluation stage re-dwells tell a similar story. The eco-focused group looked at more products and had more re-dwells at the second supermarket (in support of hypothesis H6.2 and H4.2 respectively), whereas the difference across supermarkets for the brand-focused shoppers was very small. Overall, we found an inverse correlation between preference congruence and number of products focused. The eco-focused group did end up making better decisions at supermarket 2 than in supermarket 1 whereas the brand-focused and control groups made worse decisions in the second supermarket. When comparing the groups at the second supermarket the differences in decision quality were insignificant.

At the yoghurt shelf in the second supermarket the organization is the same as their pasta shelf. Each brand has its own block even though in this case some of them have as few as two rows but many more facings per row. This makes the comparison to the pasta data somewhat problematic because the "blocks" of brand items look very different (nearly square at the pasta shelf; rectangular and long at the yoghurt shelf). The placement of the different brands' healthier options however is more in line with the eco-category at the second supermarket's pasta shelf. That is, very spread out.

When moved from the first to the second supermarket, the brand-focused group reported having an easier time choosing their yoghurt product (in support of hypothesis H1.2). A paired within-group t-test bore this out ($t(19) = -2.08$, $p = 0.05$; fig. 6 in *Results*). A paired within-group t-test on the control group, comparing their data from the two supermarkets, showed that they had a slightly *harder* time at the second supermarket (fig. 7 in *Results*). Taken together with the results for the brand-focused group, this is a strong indication that it is the shelf organization that helps make finding and choosing products easier (when preferences are congruent with the organization). This is because we should expect the control group to have no support from preference congruence at the first store and that factors such as unfamiliarity and more product facings at supermarket 2 make it harder for them there. The fact that preference congruence still wins out over these factors for the brand-focused group points to the veracity of hypothesis H1.2.

Furthermore, the between-group comparison shows that the brand-focused group at supermarket 2 had the easiest time out of the three groups (H1.1).

Looking at the correlations (fig. 5 in *Results*) it is not as clear-cut as it was with the pasta data, but it does indicate that congruence between preference and shelf organization makes shopping easier.

The health-category group from supermarket 1 spent more time looking at the yoghurt shelf—both in total (H2.2) and while evaluating options (H3.2)—in the second supermarket than they did in the first. Between-group results indicate that the health-focused group was also much slower than the congruent, brand-focused group ($t(12) = 1.93$, $p = 0.077$; H2.1) and the control group. The difference between the category group's evaluation times at supermarket 1 and 2 was also much larger than the difference between the brand group's times. While both groups spent more time evaluating their options at supermarket 2—no doubt in part a due to unfamiliarity and shelf size—the brand group, who's preferences were congruent at supermarket 2, was much faster in comparison.

Staying at the yoghurt shelf in the second supermarket, the health-focused group looked a more products (H6.1) and had more re-dwells (H4.1) than both the other groups and compared to the health-focused group in supermarket 1. This was true of both within-group and between-group comparisons and for both total (H4.1; H4.2) and evaluation stage (H5.1; H5.2) re-dwells. The brand-focused was the only group to have fewer re-dwells at the second supermarket (H4.2).

As we've seen, along with *more* dwell time and more re-dwells however, better decision making tends to follow. Again this held true, at least when compared to the brand-focused group. The within-group results did not show any significant differences for the health-focused group in decision quality after the move to the second supermarket. The brand-focused participants made worse decisions than both the category-focused ones ($t(22) = -1.50$, $p = 0.15$) and the control-group ($t(25) = -1.69$, $p = 0.10$) in the second supermarket. They also made worse decisions there (congruent) than they did at the first supermarket (incongruent).

To be clear, there are several possible reasons for dwell times being longer and for having more re-dwells in general in the second supermarket. That store has a larger shelf for both pasta and yoghurt featuring more items (336 vs 211 facings for pasta; 202 vs 161 for yoghurt). For the paired (within-group) results it is also the case that the participants are more familiar with supermarket 1 (see Lingonblad, 2015 for relevant measures of familiarity and results from the same data that is used herein) than with the second. Many of them had never visited supermarket 2 at all before. This means that any results that go against the expected consequences of those differences are a strong support for our hypotheses. There are therefore good reasons to conclude that congruence between preferences and shelf organization matters and that it makes shopping easier and faster.

So, to conclude, when the organization is congruent with the preferences of the participants, they tend to spend less time in front of the shelf before choosing their product. They also, to a somewhat lesser extent, spend less time evaluating the different options. When the shelf organization is incongruent with one's preferences the opposite appears to be true. Furthermore, for all the positives that stem from having congruent preferences, the shopper will tend to make *worse* purchase decisions. This last, interesting result is something that we did not predict and which needs more study in order to be properly explained.

Having to spend less effort on finding certain products and keep track of their attributes could perhaps free up attentional resources that could be used to make careful comparisons. This should then result in bought products being more conforming to one's preferences. It would seem that we were mistaken in thinking that more effortless shopping would support better decision making. Perhaps the time and effort saved at a shelf that suits your preferences is also time and effort lost when it comes to making good decisions. Perhaps we do not use the "freed up" attentional

resources for making better decisions at all and instead end up using less resources to make good decisions because we spend less time and effort at the shelf.

Definitions

Of central importance to this study is the division into groups based on preference. It is important however, to keep in mind that "brand" is not a category in the same sense as the "ecological option" and "healthier option" categories are. The essential distinction is that a shelf cannot be organized both by brand *and* by any of these categories since one or the other will be spread out. The choice of shelf organizations to investigate in this study is based on how shelves are commonly organized in supermarkets in general and in our two supermarkets in particular. This also explains why we did not use the same categories for pasta and yoghurt (e.g. eco-products for both). There are many categories we could have chosen but we chose the way we did based on the existing layout of the shelves (e.g. eco-pasta being blocked together at supermarket 1).

The demarcation of the groups is not completely straightforward. It is to some extent arbitrary at what values the lines are drawn, but they have nevertheless been carefully chosen. We opted to strike a balance between making the definitions as strict as possible while not making the groups too small for the results to be reliable (see table 1 in *Results*). In the same table one can see that the participants who were recruited at the second supermarket tended to be more brand-focused. This is interesting since they are also the ones who are more familiar with that store (Lingonblad, 2015). Perhaps their preferences are to some extent shaped by the way their supermarket of choice is organized?

Experiment design factors

The research in this paper was carried out in parallel with other research for which the experiment was also designed, namely the effects of familiarity with a supermarket and its assortment of products. Perhaps the results in this paper would be more accurate if the experiments were designed especially with the present investigation in mind. Such an improved experiment design would include controlled organization of shelves instead of reliance on the supermarkets' shelves being different in a quantifiable way. Such an experimental design would however not be compatible with an investigation into the effects of familiarity since none of the participants would be familiar with our particular layout.

Self-report measures of ease-of-shopping could be improved with more questions. Specifically it would have been useful to make a distinction between "easy to find" and "easy to choose" since those are to some extent quantifiable using our analysis (i.e the distinction between the orientation stage and the evaluation stage). It is however questionable if participants could be relied upon to make such a subtle distinction under the circumstances. It is interesting to take subjective measures into account and so it is a drawback that we did not have participants elaborate.

We chose to visit the supermarkets one month apart. One reason for this is that people get their salaries once a month and so will have roughly the same economic situation on average. Another reason is that we wanted the participants from the first supermarket to have time to consume the bought products so that they would not lack motivation to buy the same product types again. This also limited the choice of product together with the fact that we wanted products that people don't always buy the same kind of. There are important differences between the product types that could have affected our results. For instance, people buy yoghurt more often than they buy pasta.

One factor that we could not eliminate is the fact that the returning participants were in fact planning on making purchases in the first supermarket (we recruited them at the entrance) whereas they were at the second supermarket mainly to take part in our experiment. It is reasonable to assume that they had planned to also make other purchases since they were going to the supermarket, but it is worth noting that this might affect the results in some subtle way.

Another problem with using returning participants is that it was mainly a setup to measure the effects of familiarity (again, see Lingonblad, 2015 for that part of the study) which entailed that the returning participants were in general unfamiliar with the second supermarket whereas the ones recruited at the second store were more familiar with that one. This meant that when comparing for instance the brand-focused group from supermarket 1 with the brand-focused group from supermarket 2, that second group consisted of more participants unfamiliar with the layout of the store (see table 1, p. 6).

In order to keep the setting as natural as possible, there was no time limit set and we did not show them exactly what shelf to buy from. This last part was perhaps a mistake since we had to exclude quite a lot of participants in the analysis because they bought yoghurt from the lactose-free yoghurt section or pasta from a Barilla only promotional shelf.

Measurements

It is perhaps puzzling that there is no linear correlation between total dwell-time and the number of re-dwells. It stands to reason that the longer one looks at a bunch of AOIs, the more AOIs one will tend to look at (number of focusings) and return to (re-dwells). The first part is true, that longer dwell-times entail more AOI focuses. Dwell-time did not, however, correlate with number of re-dwells. The role of re-dwells in decision making is not clear-cut and so correlations or the lack thereof can be hard to explain. Re-dwells can indicate an increase in effort but they can also suggest the search for more information about something already seen.

The number of re-dwells seems to differ a lot between the pasta and yoghurt data. This may be because of different attitudes towards the different product categories. What pasta to get is often decided by the recipe of the whole meal whereas yoghurt is more often just a matter of taste. In other words, the pasta choice might be more important than the yoghurt choice.

Eye-tracking measures of cognitive effort could take into account pupil dilation, which is a common way to measure cognitive load (Holmqvist et al., 2011). But whereas this is a great idea in a laboratory, it is not at all suited for supermarket eye-tracking research. So many other factors such as uneven lighting have larger effects on pupil dilation than does cognitive load in such settings.

All dwell-times used in this study are dwells on AOIs where a dwell is >120ms long. The raw data from the eye-tracking glasses is not divided in this way and so we had to make some important choices when analyzing the data: The first choice was to exclude data where the participant was not

looking at the products on the shelf, i.e. when their gaze was not on an AOI. By doing this we ensure that search and evaluation times were (at least most likely) spent with attention directed at the products. We also had to make a choice about what a dwell is. Looking at an AOI for just a few milliseconds is not enough to comprehend what is being looked at. We therefore had to set a threshold for where it is reasonable for one dwell to be enough for comprehension. We based this decision on the reasoning in Gidlöf et al. (2013) that 120ms is what is needed to read and understand a word, so it is fitting for looking at a product as well. What is excluded by this threshold is mainly saccades of the eyes, that is, rapid movement from one point to another where whatever is in-between is not attended. We then divided the data for each shelf (pasta and yoghurt respectively) into an orientation stage and an evaluation stage. This demarcation was drawn at the first dwell on the product that was ultimately chosen (the *Natural Decision Segmentation Model* or *NDSM*, p.2 in this paper; Gidlöf et al., 2013). It is of course quite possible that this was not always where participants ended their search and begun their evaluation of different choices. An alternative to this way of dividing the stages would be to use the first re-dwell as an indicator of the start of the evaluation stage since re-dwells often indicate evaluation. This strategy has been used by Russo & Leclerc (1994) and is discussed in Gidlöf et al. (2013). It is however problematic since re-dwells can also indicate other processes such as the gathering of more information about something already seen. NDSM permits re-dwells in all stages which enables us to test both hypothesis H4 (total re-dwells) and hypothesis H5 (re-dwells during the evaluation stage).

Finally, since the preferences of the participants are of central importance to this study it should be asked if our way of measuring them was good. The fact that their choice of products had already (and very recently) been made when they filled in the survey may be a problem since the interaction between decision making and preferences is a two-way street (Simonson, 1999). Having them fill in the survey beforehand would have brought similar problems.

Regarding decision quality which was based on the self-reported preferences, is it reasonable to assume that the participants tried to find the best alternative? Even though they were reimbursed for their purchases—and surely only bought some of the items because we asked them—they were the ones to eat the food. It seems reasonable therefore to assume that they were at least going for “max reward for minimum effort”.

Future research

As was mentioned in the *Experiment design factors* section, a specialized version of this experiment would include controlled organization of the shelves. This is something that in a real supermarket would be problematic since it would probably entail a lot of extra work and potential decreases in sales for the retailer. Instead, such an experiment would perhaps be better run using stationary eye-trackers in a laboratory. This would also make it possible to use pupil dilation measurements of cognitive load since a controlled setting would make those results more reliable.

As was mentioned in the introduction, we had participants buy products from three product categories, but only data from two have been used in this paper. That is because the raw data from the cereal shelves was not semantically coded and analyzed in time for its writing. When that is done, the data could be used in the same way as the pasta and yoghurt data reported in this paper to see if it would lend more support to the hypotheses.

Conclusion

Two questions were asked at the outset: Does the way a supermarket’s shelves are organized affect the ease with which we can choose our products? Does it affect the quality of our choice? Our results show that when the shelf is organized in a way congruent with one’s preferences it makes it easier for us to choose our product. The strongest indicator of this was the participants’ self-reported effort but there were also indications from the eye-tracking data.

Furthermore, when the shopping trip is effortless we tend to make worse decisions. That is, we choose products that match our preferences poorly. Taken together with prior research that shows that effortless shopping makes for a decrease in sales—supposedly because less time is spent browsing and making impulse purchases—it makes for some interesting considerations for retailers. Do you want your customers to spend less and make worse decisions but come away feeling that they had an effortless visit? Or do you want them to have a more difficult time while spending more money and making better decisions?

Future research—with a more controlled setup—could find stronger correlations between preference congruence and effort. It could also investigate more subjective aspects of a shopper’s experience. Lastly it would be interesting to have a larger group of participants to find what shelf organization would be most helpful to the average customer. Remains only for the retailer to decide how helpful one ought to be.

Acknowledgements

I would like to thank my supervisors Annika Wallin and Kerstin Gidlöf for their helpful input and for their efforts in making this project successful. I would also like to thank Martin Lingonblad who also wrote his master’s thesis as part of the same project. We are all thankful to the personnel at the supermarkets and to ICA for being so helpful during our data collection. This study was part of a research project supported by the Hakon Swenson foundation.

References

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *The psychology of learning and motivation*, *8*, 47-89.

Ball, D. (2004). Consumer-goods firms duel for shelf space. *Wall Street Journal*, (October 22).

Bettman, J. R., Johnson, E. J., & Payne, J. W. (1990). A componential analysis of cognitive effort in choice. *Organizational behavior and human decision processes*, *45*(1), 111-139.

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of consumer research*, *25*(3), 187-217. Cooper-Martin, E. (1994). Measures of cognitive effort. *Marketing Letters*, *5*(1), 43-56.

Clement, J., Aastrup, J., & Forsberg, S. C. (2015). Decisive visual saliency and consumers' in-store decisions. *Journal of Retailing and Consumer Services*, *22*, 187-194.

Cooper-Martin, E. (1994). Measures of cognitive effort. *Marketing Letters*, *5*(1), 43-56.

Dreze, X., Hoch, S. J., & Purk, M. E. (1995). Shelf management and space elasticity. *Journal of*

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. Cognitive Neuroscience: The biology of the mind.(2014).

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. *Journal of Eye Movement Research*, *6*(1), 1-14.

Gidlöf, K., Holmqvist, K. & Wallin, A (2014). The centre might not be so shining after all – central bias from computer screen to the supermarket, *Manuscript submitted for publication* (see Gidlöf, 2014).

Gidlöf, K. (2014) *Visual attention during decision-making in natural environments* (Doctoral dissertation, Lund University)

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, *103*(4), 650.

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.

Kaplan, D., Horton, N. J., & Pruim, R. (2015). Randomization-based inference using the mosaic package.

The Keyhole. (2015, February 24). Retrieved from http://www.livsmedelsverket.se/en/livsmedel-och-innehall/text-pa-forpackning-markning/nyckelhalet/

KRAV-märket och andra miljömärkningar. (2015, January 13). Retrieved from http://www.krav.se/krav-market-och-andra-miljomarkningar

Lingonblad, M. (2015). *Lost in the Supermarket: Studying the effects of familiarity on consumer decision making* (Master's thesis). Retrieved from Lund University Publications database http://lup.lub.lu.se/student-papers/record/7446285

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.

Nordfält, J. (2007). *Marknadsföring i butik: om forskning och branschkunskap i detaljhandeln*. Liber.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press.

Roberts, W. A. (2003). 2001 New Products Conference: Take a Chance. *Prepared Foods, November*, *21*.).

Rushworth, M. F., Kolling, N., Sallet, J., & Mars, R. B. (2012). Valuation and decision-making in frontal cortex: one or many serial or parallel systems?. *Current opinion in neurobiology*, *22*(6), 946-955.

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

Shugan, S. M. (1980). The cost of thinking. *Journal of consumer Research*, 99-111.

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

Simonson, I. (1999). The effect of product assortment on buyer preferences. *Journal of Retailing*, *75*(3), 347-370.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological review*, *79*(4), 281.