

EMPIRICAL ARTICLE

Eye movements as a tool to investigate exemplar retrieval in judgments

Agnes Rosner ^{1,2}, Fabienne Brändli², and Bettina von Helversen³

¹Leibniz University Hannover, Hannover, Germany; ²University of Zurich, Zurich, Switzerland and ³University of Bremen, Bremen, Germany

Corresponding author: Agnes Rosner; Email: agnes.rosner@psychologie.uni-hannover.de

Received: 31 October 2023; **Revised:** 5 January 2024; **Accepted:** 16 January 2024

Keywords: preferences; memory; similarity; multiattribute judgments; eye movements; process tracing

Abstract

The retrieval of past instances stored in memory can guide inferential choices and judgments. Yet, little process-level evidence exists that would allow a similar conclusion for preferential judgments. Recent research suggests that eye movements can trace information search in memory. During retrieval, people gaze at spatial locations associated with relevant information, even if the information is no longer present (the so-called ‘looking-at-nothing’ behavior). We examined eye movements based on the looking-at-nothing behavior to explore memory retrieval in inferential and preferential judgments. In Experiment 1, participants assessed their preference for smoothies with different ingredients, while the other half gauged another person’s preference. In Experiment 2, all participants made preferential judgments with or without instructions to respond as consistently as possible. People looked at exemplar locations in both inferential and preferential judgments, and both with and without consistency instructions. Eye movements to similar training exemplars predicted test judgments but not eye movements to dissimilar exemplars. These results suggest that people retrieve exemplar information in preferential judgments but that retrieval processes are not the sole determinant of judgments.

1. Introduction

Imagine the following situation. You want to buy a smoothie. You have several smoothies available that vary on different ingredients (i.e., attributes), and for each option you judge, how much you would like it (i.e., criterion). How do people form such preferential judgments? People may make their judgments by weighting and integrating the reasons (i.e., attributes) that speak for each option by using a *rule* (Brehmer, 1994; Keeney et al., 1979). For instance, when evaluating the likeability of smoothies, you may think of how much you like each individual ingredient. Several lines of research suggest that judgments may also draw on previous experiences stored in memory (Gilboa & Schmeidler, 2001; Gonzalez et al., 2003; Juslin et al., 2003; Scheibehenne et al., 2015) in line with the idea that memory plays an important role in preference formation (Shadlen & Shohamy, 2016; Weber & Johnson, 2006; Weibächer et al., 2021). Thereafter, the judgment of how much one likes smoothies may also depend on the *similarity* in attribute values to previous instances (i.e., exemplars) stored in memory (Nosofsky, 2011). The evaluation of a smoothie will therefore depend on how much one liked similar smoothies in the past.

Testing the influence of past experiences on preferential judgments is challenging, as memory processes are difficult to observe and on the outcome level (i.e., choices or judgments) rule and exemplar-based accounts often make similar predictions. Disentangling them requires the careful construction of test situations (Juslin et al., 2008). In preferential judgments, in which the importance of attributes is unknown and likely varies between people, it becomes difficult to create these situations. Recent research has shown that eye movements can be studied to gain insights into how memory influences decision-making (Krefeld-Schwalb & Rosner, 2020; Pärnamets et al., 2016; Platzer et al., 2014; Renkewitz & Jahn, 2012) and the retrieval of past experiences during inferential judgments and categorizations (Rosner et al., 2022; Rosner & von Helversen, 2019; Scholz et al., 2015). In the study by Scholz et al. (2015), participants first memorized multiple pieces of information about various job candidates (exemplars). In subsequent test trials, they judged the suitability of new candidates that varied in their similarity to the previously learned exemplars on several cues. Test items were presented auditorily while participants saw only the empty rectangles of the trained exemplars on the screen. Results showed that when using an exemplar-based decision strategy, but not when using a rule-based strategy, participants fixated longer on the previous location of exemplars that resembled the new candidates (the so-called looking-at-nothing effect) than on the locations of dissimilar exemplars. Moreover, fixation durations increased as the similarity between test items and exemplars increased (the so-called similarity effect on eye movements).

Utilizing the looking at nothing behavior, we aim to investigate memory retrieval in inferential and preferential judgments and to gain further knowledge on the role of retrieval processes within the judgment process. In the following, we first review the literature about the processes involved in multiple cue judgments and why exemplar retrieval might occur in preferential judgments. Next, we outline how we measured retrieval processes during inferential and preferential judgments, discuss how eye movements can contribute to understanding the role of memory retrieval in judgment, and present 2 experiments to test our hypotheses.

1.1. Exemplar-based processes in inferential and preferential judgments from multiple cues

There is a large amount of research on strategy use in inferential judgments (see Pachur & Bröder, 2013 for an overview). Thereafter, people preferentially form judgments from multiple cues by using linear additive rules (Bröder & Gräf, 2018; Juslin et al., 2008). Accordingly, people abstract the cue-criterion relation for each cue, weigh each cue by its importance, and form a judgment by summing up the weighted cue values (e.g., Brehmer, 1994; Juslin et al., 2003). Usually, these rule-based cognitive processes are assumed to be of a reflective and deliberate nature and depend on working memory capacity (Hoffmann et al., 2014; Juslin et al., 2008).

Juslin et al. (2003, 2008) introduced the retrieval of exemplars as an alternative process to the judgment literature. Exemplar-based models assume that people's judgments and decisions are influenced by the similarity of options to previously encountered instances that are stored in memory (Hoffmann et al., 2016; Medin & Schaffer, 1978; Nosofsky, 2011; von Helversen & Rieskamp, 2009). Thereafter, people retrieve patterns of cue values and their relations to a judgment criterion from memory. The exemplar with the highest similarity (e.g., the number of matches in cue values) determines the judgment. Exemplar-based processes are often assumed to be implicit, automatic processes (Hahn et al., 2010) that rely on episodic memory (Hoffmann et al., 2014).

The introduction of exemplar-based processes to the judgment literature has spurred investigation into when they are more likely used. For instance, in environments, in which cue-criterion relations are hard to extract, people shift toward an exemplar-based strategy (Bröder et al., 2010; Platzer & Bröder, 2013; von Helversen et al., 2013). Furthermore, an environment that consists of a multiplicative combination of cues can foster the use of an exemplar-based judgment strategy (Hoffmann et al., 2016; Juslin et al., 2003). Next to cue complexity, learning, and memory also influence the use of an exemplar strategy. Bröder and Gräf (2018) demonstrated that having to retrieve cue information from memory instead of presenting cue information on screen adds to the effect of cue complexity on exemplar-based

strategy use. Moreover, a learning environment in which people directly learn about each training item (instead of through paired comparisons) increases the use of an exemplar-based strategy (Pachur & Olsson, 2012). In sum, it seems that people often prefer using rules but can adaptively switch to an exemplar-based strategy when the environment calls for a strategy shift (Bröder & Gräf, 2018; Juslin et al., 2008; Karlsson et al., 2008).

According to the ‘division of labor’ hypothesis (i.e., Juslin et al., 2008), people either rely on an exemplar-based strategy *or* a rule-based strategy. However, further research proposes that judgments may be formed by a ‘blending’ or mixture of both exemplar-based and rule-based processes *within* a single judgment (Albrecht et al., 2020; Bröder et al., 2017; Erickson & Kruschke, 1998; Herzog & von Helversen, 2018; Hoffmann et al., 2014; Izydorzyc & Bröder, 2021; Schlegelmilch et al., 2022; von Helversen et al., 2014). Thereafter, people can adaptively exploit both their memories of previous experiences and combine information following a rule-like integration process. This view has been stipulated by findings that show that judgments about job applicants can be biased by facial similarity to previously encountered job candidates, even when information about the applicants’ attributes such as their work experiences or abilities is relied upon in a rule-based manner (von Helversen et al., 2014). In line with the assumption of both rule and exemplar processes determining single judgments, the CX-COM model by Albrecht et al. (2020) assumes that exemplars in memory compete for retrieval and one exemplar is recalled from memory. In the second step, the retrieved judgment is adjusted based on abstracted cue knowledge. Blending models like the measurement model RuleEx-J for judgments (Bröder et al., 2017; Izydorzyc & Bröder, 2021) assume that an exemplar- and a rule-based judgment process act in parallel. The resulting judgment is a weighted combination of both interim judgments. Neural evidence supports the notion of several processes contributing to judgments at the same time. Wirebring et al. (2018) demonstrated that both processes lead to the activation of similar brain regions. In sum, according to mixture models, judgments are both influenced by the retrieval of exemplars and the integration of cue information in a rule-based manner.

Most research on relying on exemplar memory during judgment and decision-making has focused on inferential judgments. Inferential decisions reflect the decision-maker’s expectations concerning the consequences of an uncertain outcome. These expectations can be compared to objective consequences. For instance, if selecting a smoothie not for oneself, but for a friend, one can assess whether the chosen smoothie was indeed to the friend’s liking. In other words, in inferential choices, there is an objective criterion outside of the decision-maker that can be used to decide whether the choice was correct or incorrect. However, many common real-life judgments and decisions are based on preferences. For instance, many people will choose a smoothie based on their own subjective utilities. In this preferential decision situations, there is no objectively correct or incorrect decision option outside of the decision-maker, but only options that align with the decision-maker’s preferences.

Little direct evidence exists that people retrieve exemplars from memory when making preferential judgments. There are several lines of research and theoretical accounts arguing that exemplar processes may also play a role in preferential judgments. In case-based decision-making (Gilboa & Schmeidler, 2001), it is assumed that if a decision-maker must decide without knowing the task structure and does not receive feedback after the decision, a similarity-based decision process becomes more likely (Grosskopf et al., 2015; Ossadnik et al., 2013). Building on this work, a recent model by Bordalo et al. (2020) assumes choices result from choice options cueing the recall of similar past experiences which form a norm that then serves as an initial anchor for valuation.

In instance-based learning, current decisions depend on previous decisions in a dynamic decision context (Gonzalez et al., 2003). That is, during the retrieval of past instances stored in memory, preferences for current decisions are formed.

In consumer choices, it has been shown that similarity to well-known and successful brands can increase the choice share of a copycat product (van Horen & Pieters, 2012; Warlop & Alba, 2004). Furthermore, similarity influences how people categorize new products and how these categorizations affect brand attitudes (Basu, 1993; Lajos et al., 2009) and price estimates (Scheibehenne et al., 2015).

Last, even hedonic judgments seem to be influenced by the similarity to previously encountered situations (Martindale, 1984). That is, the more strongly a similar esthetic experience (e.g., a sunset) is reactivated from memory, the stronger its influence on the current situation. In a similar vein, the evaluation of novel faces is influenced by their similarity to familiar faces (Verosky & Todorov, 2010).

In support of the idea that exemplars influence preferential choices, Jarecki and Rieskamp (2022) studied an incentivized multi-attribute choice task. In this task, participants initially stated their willingness to pay for 8 out of 16 choice options (pens varying on 4 attributes). Next, participants completed a 2-alternative choice task in which they repeatedly chose between all pens and a monetary value that was based on their initial and individually varying willingness to pay evaluations. Using cognitive modeling of choice strategies, they showed that most participants were better described by a memory-based value model considering the similarity to previous instances stored in memory rather than a multi-attribute value model relying on linear additive integration.

There are also reasons why in preferential judgments exemplar retrieval may play less of a role than in inferences. Preferences may rather be driven by habits (Aarts et al., 1998; Verplanken & Orbell, 2022). Thereafter, preferential judgments would be based on learning experiences from repeatedly occurring past events. For instance, choosing a smoothie every morning before work. Habitual behavior does not afford considering and integrating attribute values, because stable preferences were already built. Still, even habitual behavior will entail some form of episodic memory retrieval. Exemplar retrieval may not be as important in preferential judgments as immediate emotional reactions to the options may guide these judgments instead (Loewenstein et al., 2001; Rozin & Todd, 2015; Zajonc, 1980, 2000). In such a case, the judgment process may nonetheless involve retrieval from episodic long-term memory, for instance, the retrieval of emotional experiences (Betsch et al., 2001).

In sum, recent research suggests that exemplar retrieval might replace or occur in addition to rule-based processes during preferential judgments, but on the process level, little direct evidence has been reported. Here, we aimed to use eye tracking to be able to compare process-level data related to memory retrieval in general and particularly in exemplar retrieval in preferential compared to inferential judgments tasks. In the following, we outline why eye tracking can be a suitable tool for the job.

1.2. Tracing retrieval processes through eye movements

Eye movements are quick, frequent, and highly automatic actions (Rayner, 2009) that can reflect attention and information search in a variety of tasks, including judgment and decision-making (Orquin & Mueller Loose, 2013; Schoemann et al., 2019; Schulte-Mecklenbeck, Johnson, et al., 2017; Schulte-Mecklenbeck, Kühberger, et al., 2017).

Eye movements can also be used to trace memory processes (Peterson & Beck, 2011). When people retrieve information from memory, they look at spatial locations where the information was originally presented—even if the information is no longer visible (Altmann, 2004; Ferreira et al., 2008; Johansson et al., 2006; Laeng et al., 2014; Martarelli & Mast, 2013; Richardson & Kirkham, 2004; Richardson & Spivey, 2000; Scholz et al., 2016; Wantz et al., 2016; Wynn et al., 2019). In the classic paradigm, Richardson and Spivey (2000) presented participants with a spinning cross in 1 of 4 equal-sized areas on a computer screen together with spoken factual information. In a later test phase, participants heard a statement regarding the presented facts and had to judge the truth of the statement. Even though during this retrieval phase the computer screen was blank, participants fixated more often on the spatial area where the sought-after information had been presented than on the other 3 areas on the screen.

Most likely, people show this *looking-at-nothing-effect* because during encoding, information from multiple sources of input, including the locations of perceived objects, is integrated into an episodic memory representation. Once the episodic memory representation is reactivated during retrieval, it spreads activation to the motor system (via feeding in activations in a shared priority map (Awh et al., 2012; Hedge et al., 2015; Theeuwes, 2018), which in turn leads to the execution of eye movements back to the locations linked with the memory representation (Huettig et al., 2012; Wynn et al., 2019).

Researchers have used this phenomenon to trace what information is retrieved from memory during a wide range of JDM tasks (Jahn & Braatz, 2014; Klichowicz et al., 2021; Krefeld-Schwalb & Rosner, 2020; Pärnamets et al., 2016; Renkewitz & Jahn, 2012; Rosner et al., 2022; Rosner & von Helversen, 2019; Scholz et al., 2015, 2017). In the study by Renkewitz and Jahn (2012), participants first learned attribute information about decision options that were arranged within spatial frames at distinct spatial locations on a computer screen. Later they were asked to choose between the options. During decision-making, eye movements to the emptied spatial locations reflected information search in memory for the learned attribute information. This study demonstrates that eye movements based on the looking-at-nothing behavior can reveal what information is activated in memory during JDM.

Rosner and von Helversen (2019) studied looking-at-nothing in inferential judgments. They established a direct link between looking-at-nothing and the resulting behavioral judgment about the suitability of job candidates. The more participants looked at high-performing exemplars, the more their judgments resembled those of the high-performing training exemplars. These results suggest that exemplar-based retrieval processes occurring during judgment and decision-making can be made visible by recording eye movements at the previous locations of the exemplar information.

1.3. The present research

We aimed to use looking-at-nothing behavior to investigate whether we find evidence for exemplar retrieval in preferential judgments and to link the looking-at-nothing behavior to the judgments. We used a multiple-cue judgment task, in which participants learned about 4 training exemplars at different locations on the screen during a training phase. We then recorded eye movements to the previous locations of the training exemplars while they judged test items that differed in their similarity to the training exemplars. We compared eye movements during a preferential judgment task, in which participants judged their preference for different smoothies based on their ingredients, with eye movements during an inferential task, in which participants had to infer the smoothie preferences of another person based on the training exemplars.

1.4. Research questions

As outlined above, the looking-at-nothing phenomenon has been shown to be a reliable indicator of retrieval processes. Thus, a lack of evidence of eye movements to the locations of the training exemplars in preferential compared to inferential judgments would suggest less reliance on similar past experiences in preferential judgments. In contrast, eye movements to the location would suggest that participants are indeed retrieving information about the training instances during the test.

However, even though evidence for the retrieval of exemplar information provides the necessary foundation for the involvement of exemplar-based processes in the judgment process and the study by Scholz et al. (2015) suggests that rule-based processing by itself does not induce looking-at-nothing behavior, it does not necessarily mean that exemplar-based processes are governing the judgment process. Indeed, there are 3 broad possibilities for how retrieved information could be used in the judgment process: (1) people rely predominantly on the retrieved exemplar information during judgment, (2) people retrieve information about the exemplars but do not solely rely on retrieved exemplars but also on further processes such as rules or preexisting attitudes, and (3) people retrieve information, but do not rely on this information during judgment.

If people predominantly rely on retrieved exemplars, we should find strong correlations between the judgment of a test item and the rating of the training exemplar they looked at most during the trial. In addition, the correlation should be stronger the more people look to a single training exemplar. In contrast, the more people rely on other processes such as rule-based knowledge, the less the relation between ratings of training exemplars and test judgments would depend on eye movements to the exemplar locations and the less eye movements to exemplar locations would predict test judgments beyond judgments based on rules.

We conducted a first experiment to test whether eye movements to the previous locations of training exemplars also occur in a preferential judgment task and to investigate the link between eye movements to the exemplars and judgments during the test phase.

2. Experiment 1

Participants' task was to rate smoothies consisting of different ingredients on a 7-point Likert scale. Half the participants judged 4 exemplar smoothies (training phase) first and then, in the test phase, a test set of smoothies based on their own preferences (preference condition). The other half learned the preferences of a previous participant in the preference condition for 4 exemplar smoothies (training phase) and then had to infer the preferences of this participant for the test set of smoothies in the test phase (inference condition).

All experimental materials, data, analysis scripts, and preregistration for Experiment 1 can be found on OSF (<https://osf.io/fjmkd/>).

2.1. Method

2.1.1. Participants

Previous studies found large effects (e.g., $\eta_p^2 \approx .29$) regarding the influence of similarity on eye movements (Scholz et al., 2015). To be able to find a small to medium effect of the within-subject's factor similarity with 4 levels and 12 repeated measures per similarity level, a sample of 36 participants per condition would have the appropriate power ($\eta_p^2 = .03$ requires $n = 36$ per condition to reach a power of 95%; Faul et al., 2007). Because eye-movement measures can be noisy (e.g., drop-off in data accuracy, inability to track a participant's eye movements), we aimed at testing 40 participants per condition.

Overall, 84 participants at the University of Zurich took part in the study for course credit or financial compensation (13 Swiss francs [CHF] per hour) and a voucher for one smoothie in a smoothie shop. One participant did not follow the instructions, as became clear in a post-experiment interview; for 1 participant, responses deviated more than 3 *SD* from the sample mean, and for 1 participant, the eye tracker could not be calibrated to a sufficient accuracy ($<1.5^\circ$ of visual angle). Therefore, these participants were excluded. In total, we could analyze the data of 81 participants (54 female, $M_{\text{age}} = 27.3$ years, range 19–57 years). All participants had normal or corrected-to-normal vision. Mean tracking accuracy in the test trials was very high at $M = 0.7^\circ$ of visual angle. All participants signed informed consent forms. Forty participants were in the preference condition and 41 in the inference condition. During participant recruitment, the first 40 participants were assigned to the preference condition and the second to the inference condition, because the task of the inference condition was based on the ratings from participants of the preference condition.

2.1.2. Apparatus

Participants were seated in front of a 22-inch computer screen ($1,680 \times 1,050$ pixels) at 700 mm and instructed to position their heads in a chin rest. The eye tracker system SMI iView RED sampled data from the right eye at 500 Hz and recorded with iView X 2.8 following a 5-point calibration. Fixation detection was done with IDF Event Detector 9 (SMI, Teltow) using a peak velocity threshold of $30^\circ/\text{s}$ and a minimum fixation duration of 80 ms.

2.1.3. Materials

The study materials consisted of 4 training exemplars and 36 additional test items (test smoothies). All items contained information on 4 ingredients (the cue values, e.g., apple juice, bananas, beets, raspberries). Test items varied in their *similarity* to the exemplars, ranging from 1 to 4 matches in cue values with 1 exemplar, resulting in 12 items for each similarity level. Items with one match shared one cue value with each of the 4 exemplars and were therefore ambiguous. Items with 2 matches shared 2

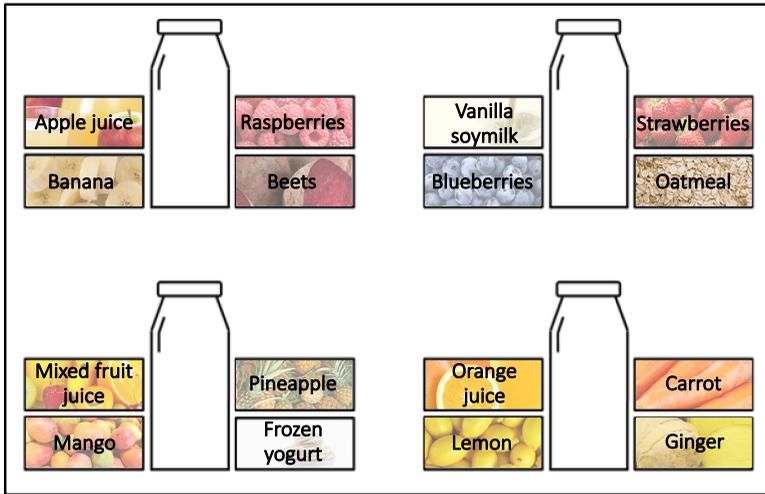


Figure 1. The 4 training exemplars (smoothies) with ingredients. Each exemplar was presented in 1 of the 4 screen quadrants. Note that the size of the exemplar on the screen is increased and colors were adapted to increase readability. See the online article for the color version of this figure.

values with 1 exemplar and 2 additional cue values with 2 other exemplars. The 4 training exemplars were included in the test set, and each was repeated 3 times during the test phase. Test materials were fully balanced. That is, each cue value and each combination of 2 cue values occurred similarly often. Consequently, each training exemplar was equally often the one with the highest number of matches (see Appendix A for a full list of items).

Visual materials consisted of the 4 training exemplars (Figure 1). Each exemplar was presented in 1 of the 4 screen quadrants. The distance from the center of the screen to the center of each of the 4 exemplars was 9.05° of visual angle (542 pixels). Cue values were presented as black text in rectangles with white borders and backgrounds consisting of photos of the ingredients. The screen background was light gray. Each rectangle had a size of $2.84^\circ \times 1.17^\circ$ of visual angle (170×70 pixels). Visual materials were presented in 4 balanced orders, varying the positions of the exemplars on the screen and the order of cue values between participants. During the subsequent test phase, participants saw only the empty rectangles on the screen. Stimuli were read out loud from left to right and from top to bottom, following the visual presentation of exemplars during the training phase.

2.1.4. Procedure

At the beginning of the experiment, the eye tracker was calibrated to test if the participants could be tracked with sufficient accuracy. Responses in the training and judgment phases were given on a Likert scale. Before the beginning of the training phase, participants practiced clicking on a particular position of the rating scale within 5 seconds. The position to click on ranged from 1 to 7 and was randomly selected and presented before each practice trial. Rating scale practice ended when participants hit the correct position 5 times.

2.1.4.1. Training phase

One trial of the training phase consisted of the visual presentation of one exemplar smoothie. Participants in the preference condition could first study the exemplar for as long as they wanted. After a mouse click, a 7-point Likert scale appeared below the exemplar smoothie (Figure 2). They had to indicate how much they would like the smoothie. There was no time restriction on the preference rating. In one round, participants rated all 4 exemplar smoothies. In total, they worked through 4 rounds.

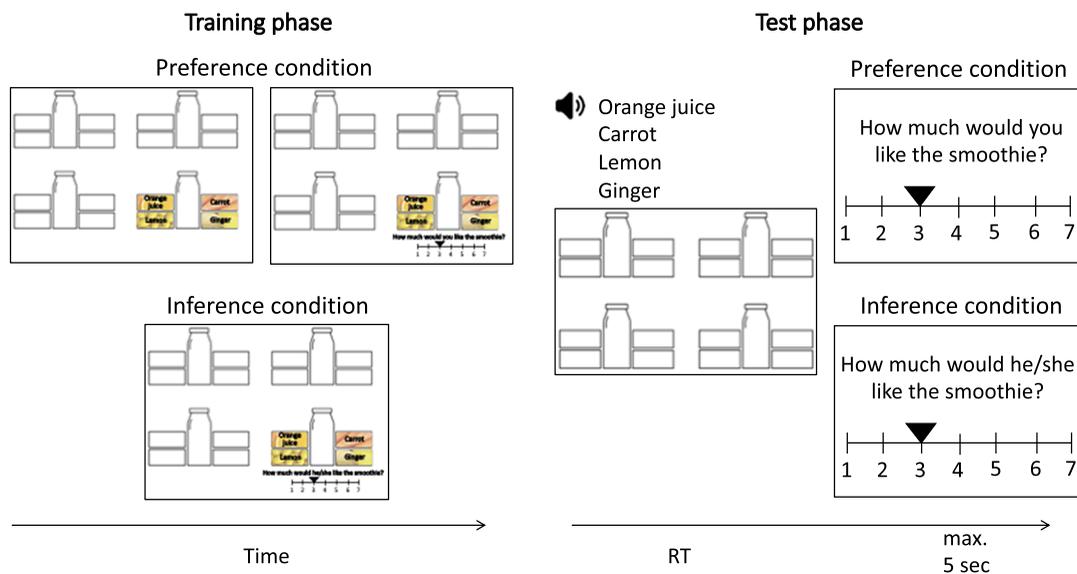


Figure 2. Procedures of the training (left) and test (right) phases of Experiment 1. See the text for a detailed description. See the online article for the color version of this figure.

Participants were instructed to respond as consistently as possible to give a representative impression of their preferences.

Participants in the inference condition saw the same exemplar smoothies. However, each exemplar smoothie was immediately presented together with the rating scale. They could not choose a value. Instead, the fourth rating of one of the participants of the preference condition was presented on the rating scale (Figure 2). Participants' task was to learn the ratings. Mouse clicks ended the presentation of one exemplar and started the presentation of another exemplar. Each participant in the inference condition was randomly assigned to the preference judgments of one participant in the preference condition. This way, we made sure that the distribution of ratings was the same in both experimental conditions.

2.1.4.2. Test phase

At the beginning of each trial, participants were asked if they were ready to start the trial. Next, a fixation cross was presented in the center of the screen for 1.5 s. Participants then listened to the cue values of 1 of 40 test smoothies while the screen contained only empty rectangles (Figure 2). Overall, participants judged 48 smoothies as smoothies with the same cue pattern as the training exemplars were presented 3 times. With a click, participants proceeded to the rating scale, which was presented in the center of the screen. We presented the rating scale on a separate screen to avoid additional eye movements driven by a visual search of the screen and the preparation of the motor response (i.e., entering the rating on the scale). Participants in the preference condition had to indicate how much they would like to drink the presented smoothie. They were instructed to respond as consistently as possible with the ratings they provided during the training phase. Participants in the inference condition had to infer how much the participant in the preference condition liked the smoothie based on the information they had learned in the training phase. Test items were presented in randomized order. Eye movements were recorded throughout the test phase.

2.1.4.3. Location and preference memory tests

All participants were asked to remember the cue values of the training smoothies. For this, as in the test phase, they saw the screen with the empty rectangles of the 4 training smoothies. Cue values were

presented auditorily and in randomized order. After hearing 1 cue value, participants had to click on the rectangle where they thought the cue value had been presented during training. The memory test ended after participants responded to all 16 cue values once.

Only participants in the inference condition were additionally tested on their memories of the learned preference ratings for the 4 exemplar smoothies. For this, they saw the cue values of 1 exemplar smoothie together with the rating scale below the exemplar, similar to the presentation in the training phase. They had to click on the position of the rating scale corresponding to the preference rating they had learned during the training phase.

2.1.4.4. Ingredient-rating test

All participants had to rate each ingredient. For this, participants were visually and auditorily presented with 1 ingredient displayed in the center of the screen, together with the rating scale. Participants in the preference condition had to indicate how much they liked the presented ingredient. Participants in the inference condition had to indicate how much the participant in the preference condition might have liked the ingredient.

At the end of the experiment, participants were asked how much they liked smoothies in general on a scale of 0 (*not at all*) to 4 (*very much*), how often they had drunk a smoothie within the last month (0 = 0 times, 1 = 1–3 times, 2 = 4–6 times, 3 = more than 6 times), and if they were allergic to any of the smoothie ingredients used in this study. The experiment lasted on average 40 min.

2.2. Results

The aim of Experiment 1 was to test if the similarity effect observed in the looking-at-nothing behavior (Rosner & von Helversen, 2019; Scholz et al., 2015) occurs for preferential judgments and to link preferential judgments to eye movements. To achieve this goal, we compared the memory-driven eye-movement behavior in a preference and an inference condition. Before reporting the results of the gaze analyses, we provide an overview of the behavioral measures testing participants' performance in the tasks and their preference for smoothies in general.

2.2.1. Preparatory data analyses and rationale for the analyses

For the analyses of behavioral measures, we aggregated data on the participant level. We analyzed the data with linear regressions with the behavioral measure (e.g., location memory performance) as a dependent variable and task (inference, preference) as an independent variable. For the eye-movement analyses, we used a mixed-model approach. Analyses were performed with R (R Core Team, 2021) and the following packages: lme4 (Bates et al., 2015), afex (Singmann et al., 2022), and estimated marginal means with emmeans (Lenth, 2022).

To analyze eye movements, we drew 4 rectangular areas of interest (AOIs) around each of the 4 smoothies shown in Figure 1. Each exemplar AOI had a size of $11.03^\circ \times 5.68^\circ$ of visual angle (660×340 pixels). The size of the exemplar AOI exceeded the outer borders of each of the 4 rectangles describing 1 exemplar by half the width of each rectangle (85 pixels).

Poisson-distributed dependent variables (e.g., count of the number of fixations on the matching exemplar location) were analyzed with a generalized linear mixed model (GLMM) analysis with a logistic link function and Laplace approximation of parameter values. The *p* values were estimated with the likelihood ratio test. We analyzed normal-distributed dependent variables (e.g., test ratings) with linear mixed models (LMM) fit by REML. *P* values were estimated using Satterthwaite's method. All models used sum of squares contrast coding. We aimed at implementing the maximal random effect's structure justified by the design (Barr et al., 2013). Continuous variables were centered on 0 (e.g., consistency, location memory performance measures, rating of exemplar looked at most).

Table 1. Results of behavioral measures in Experiment 1.

Variable	Condition				<i>b</i>	<i>t</i> (79)	<i>p</i>	<i>R</i> ²
	Preference		Inference					
	<i>M</i> (<i>SD</i>)	95% CI	<i>M</i> (<i>SD</i>)	95% CI				
Learn–test consistency	0.58 (0.33)	[0.5, 0.7]	0.36 (0.40)	[0.2, 0.5]	–0.22	–2.67	.009	.07
Location memory test								
Exemplar level	11.5 (3.11)	[10.4, 12.6]	12.7 (3.80)	[11.6, 13.8]	1.23	1.59	.115	0.2
Ingredients level	6.7 (3.92)	[5.4, 8.0]	8.80 (4.29)	[7.5, 10.1]	2.11	2.30	.024	.05

Note: Learn–test consistency denotes the mean absolute deviation between the last rating of the exemplar in the training phase and the 3 ratings collected per exemplar in the test phase. The location memory test performance denotes the number of ingredients participants correctly localized on the *exemplar level*, that is, the ingredient was assigned to the correct exemplar location (chance level 1 out of 4), and the *ingredients level*, that is, the ingredient was assigned to the correct ingredient location (chance level 1 out of 16), [Figure 1](#); CI, confidence interval.

2.2.2. Behavioral measures

2.2.2.1. Postquestionnaire

Participants did not report any serious food allergies concerning the ingredients used in the study. All participants liked smoothies (all values larger than 1) and a large part of the sample regularly drank smoothies at least once per month (0 times: $n = 24$, 1–3 times: $n = 36$, 4–6 times: $n = 14$, more than 6 times: $n = 7$).

2.2.2.2. Preference judgments in the preference condition

We aimed for a wide distribution of ratings in the preference condition to be able to test for the accuracy of judgments in the inference condition. Indeed, preference judgments ranged over the whole rating scale—both for exemplars, $M_{Ex} = 4.6$ ($SD_{Ex} = 0.6$), $range_{Ex} = 1–7$, 95% confidence interval (CI) [4.4, 4.8], and for single ingredients, $M_{In} = 4.7$ ($SD_{In} = 0.5$), $range_{In} = 1–7$, 95% CI [4.6, 4.9]. See [Appendix B](#) for an overview on the level of single exemplars and ingredients.

2.2.2.3. Inference accuracy, location, and preference memory performance

For participants in the inference condition, we tested how well their test judgments overlapped with the judgments of the matching preference participant. Participants were accurate in their inferences ([Supplementary Material](#)). Additionally, participants in the inference condition had to remember the 4 preference ratings they had learned in the training phase. On average, participants remembered 3.76 of 4 preference judgments ($SD = 0.54$), that is, 94% of all ratings.

Location memory performance describes the number of ingredients that were correctly localized either on the level of ingredients or on the level of exemplars. For the analyses on the ingredients level, the response was correct when participants selected the rectangle in which the ingredient had been shown (i.e., chance level was 1 in 16). On the exemplar level, the response was counted as correct, when the participant selected 1 of the 4 rectangles belonging to the exemplar, in which the ingredient had been shown (i.e., chance level was 1 in 4). Participants in both conditions had good location memories, but location memory was higher when analyzing it on the level of exemplars rather than single ingredients¹ ([Table 1](#)).

2.2.2.4. Consistency

We aimed to test to what extent participants followed the instructions and used the ratings they provided during the training phase for the test judgments. We calculated the mean absolute deviation between the last rating of the exemplar in the training phase and the 3 ratings collected per exemplar in the test

¹This also holds when controlling for the guessing rate by subtracting wrong responses divided by the number of response options (exemplar level: 4, ingredients level: 16) minus one (e.g., Budescu & Bar-Hillel, 1993).

phase. Participants in the preference condition were less consistent than participants in the inference condition (Table 1). Note that smaller values mean higher consistency.²

2.2.2.5. Interim summary

Overall, participants performed well in the task. They built memories of the trained exemplars and were consistent in their judgments. However, conditions slightly differed in the consistency measures.

In the following gaze analyses, we included consistency as well as location memory as covariates to rule out alternative hypotheses. For instance, as looking-at-nothing reflects bindings between spatial locations on the screen and semantic information held in memory (Wynn et al., 2019), the better participants' location memory, the more looking-at-nothing they might show.

2.2.3. Similarity effect on eye movements

Previous research has shown that the higher the similarity between test items and exemplars, the more participants will look at the associated but empty exemplar locations for inferential choices and judgments (Rosner & von Helversen, 2019; Scholz et al., 2015). We tested if this result can be replicated in the inference condition in this study and if the same effect can be found for preferential judgments.

We analyzed fixation proportions based on the number of fixations. That is, for each trial, we determined the number of fixations³ on the most similar exemplar location and divided it by the summed number of fixations on all 4 exemplar locations. Note that a similarity of 1 means the item shared 1 cue value with each exemplar. Which exemplar was considered the most similar for items with similarity = 1 was randomly determined before the beginning of the data collection. We ran a GLMMs for the binomially distributed dependent variable fixation proportion. We added task condition and similarity as well as their interaction as fixed effects. The random effect's structure consisted of by-subject random intercepts and by-subject random slopes for exemplars and similarity. We ran 2 models, one including learn-test consistency and one including location memory performance as covariates.⁴

In line with previous research, we found that with increasing similarity between test items and exemplars, participants looked proportionally more at the most similar exemplar location, $\chi^2(3) = 57.62, p < .001$ (Figure 3). But we also observed a main effect of task condition, $\chi^2(1) = 7.58, p = .006$. The increase in fixation proportions over the levels of similarity was stronger for the inference condition in comparison to the preference condition, as shown by an interaction of task and similarity, $\chi^2(3) = 10.83, p = .013$.

Contrast analyses with Holm correction revealed significant differences between all levels of similarity in the inference condition (all $ps < .001$). In the preference condition, differences between levels of similarity occurred between the first 2 and last 2 levels of similarity (all $ps < .01$), but not between similarities 1 and 2 or between similarities 3 and 4 (all $ps = 1.0$). Comparing levels of similarity between task conditions revealed that participants in the inference condition looked proportionally more at the most similar exemplar location at levels 2 and 4 of the factor similarity (all $ps < .023$), but not at the first and third level (all $ps > .10$). We observed a main effect for learn–test consistency, $\chi^2(1) = 7.09, p = .008$. The more consistent the ratings, the more people looked at the most similar exemplar location ($z = -2.73, p = .006$). That was the case in both task conditions (all $ps > .07$). There were no interactions with any of the other predictors (all $ps > .17$).

Adding the number of correct retrievals in the location memory test on the exemplar level as a covariate showed that the better the location memories, the higher the proportion of fixations on the relevant exemplar location, the main effect of number correct, $\chi^2(3) = 22.21, p < .001$ (linear contrast: $z = 5.08, p < .001$). This effect was especially pronounced for similarity levels 2 to 4 (all $ps < .001$),

²We also calculated the consistency between repeated judgments within the test phase. Results can be found in the [Supplementary Material](#).

³In comparison to the preregistration, for model fitting purposes, we based our analyses on fixation proportions calculated from the number of fixations rather than the fixations durations. Both measures correlate highly: $r(81) = .99$.

⁴Running the models on the similarity effect on eye movements without covariates revealed the same results pattern. The same is true when running the model without items being identical to a training item (similarity = 4).

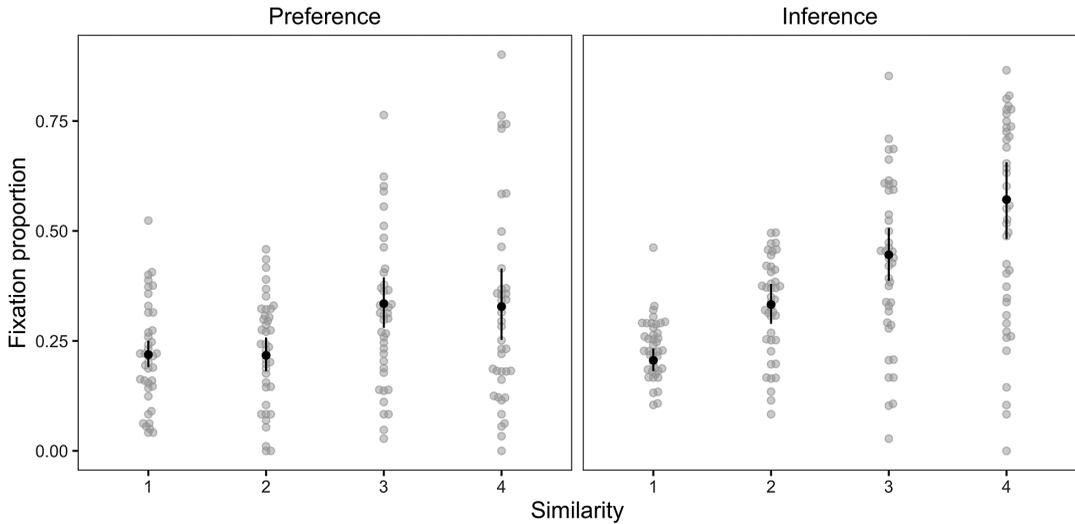


Figure 3. Estimated means for proportions of fixation on the most similar exemplar location as a function of the similarity of test items to exemplars for the 2 task conditions. Standard errors show estimated within-subject 95% confidence intervals. Gray dots show individual participants' means.

but not for similarity level 1 ($p = .31$), resulting in an interaction of similarity and location memory test performance, $\chi^2(3) = 27.72$, $p < .001$. All other results stayed the same.

In sum, the similarity effect also occurred for preferential judgments, indicating that people retrieved the training exemplars during judgment. However, the effect was reduced compared to inferential judgments.⁵

2.2.4. Eye movements and judgments

To investigate the link between the looking-at-nothing behavior and judgments—and thus which role the retrieval of exemplar information played during the judgment process—more closely, we conducted 2 further exploratory analyses. These analyses were not part of the preregistration as they were encouraged by the reviewers of this manuscript.

2.2.4.1. Looking at nothing moderates judgments

Firstly, we tested whether the relation between training and test rating is moderated by the amount of eye movements at the location of the most similar exemplar during the test. We ran a linear mixed model for test ratings with predictors for training rating of the most similar exemplar, task condition, and fixation proportions to the most similar exemplar and their interactions as well as random intercepts for subjects and exemplars. The variables training rating and fixation proportions were centered on 0. We found a main effect of training rating (Table 2). That is, preferential judgments for the most similar training exemplars predicted judgments during the test. The analyses further revealed a significant 2-way interaction between the training rating and fixation proportions.

That is, the relation between training and test ratings was moderated by the amount of looking-at-nothing. We also found a 3-way interaction of training rating, fixation proportions, and task condition indicating that the moderating effect was more pronounced in the inference as compared to the preference condition (Figure 4A). As a control, when running the same model, but adding fixation proportions to a randomly chosen exemplar as a predictor, the moderating influence was not significant anymore, interaction training rating and fixation proportions: $F(1, 3839.71) = 0$, $p = .978$ (Figure 4B).

⁵An analysis of the looking-at-nothing behavior for items with identical cue patterns as the training candidates can be found in the [Supplementary Material](#).

Table 2. Result of the mixed model analyses on test judgments of Experiment 1.

Fixed effect	<i>b</i>	<i>SE</i>	<i>F</i>	<i>df</i>	<i>p</i>
(Intercept)	4.22	0.04			
Training rating	0.70	0.02	1005.76	1, 365.37	<.001
Fixation proportions	0.07	0.02	10.13	1, 3796.42	.001
Task condition	-0.13	0.04	9.98	1, 78.14	.002
Training rating × Fixation proportions	0.22	0.02	116.57	1, 3878.31	<.001
Training rating × Task condition	0.05	0.02	5.18	1, 3874.79	.023
Fixation proportions × Task condition	0	0.02	0	1, 3824.07	.974
Training rating × Fixation proportions × Task condition	-0.09	0.02	16.66	1, 3877.99	<.001

Note: CI, confidence interval. Task conditions were coded as 1 = Preference, 2 = Inference.

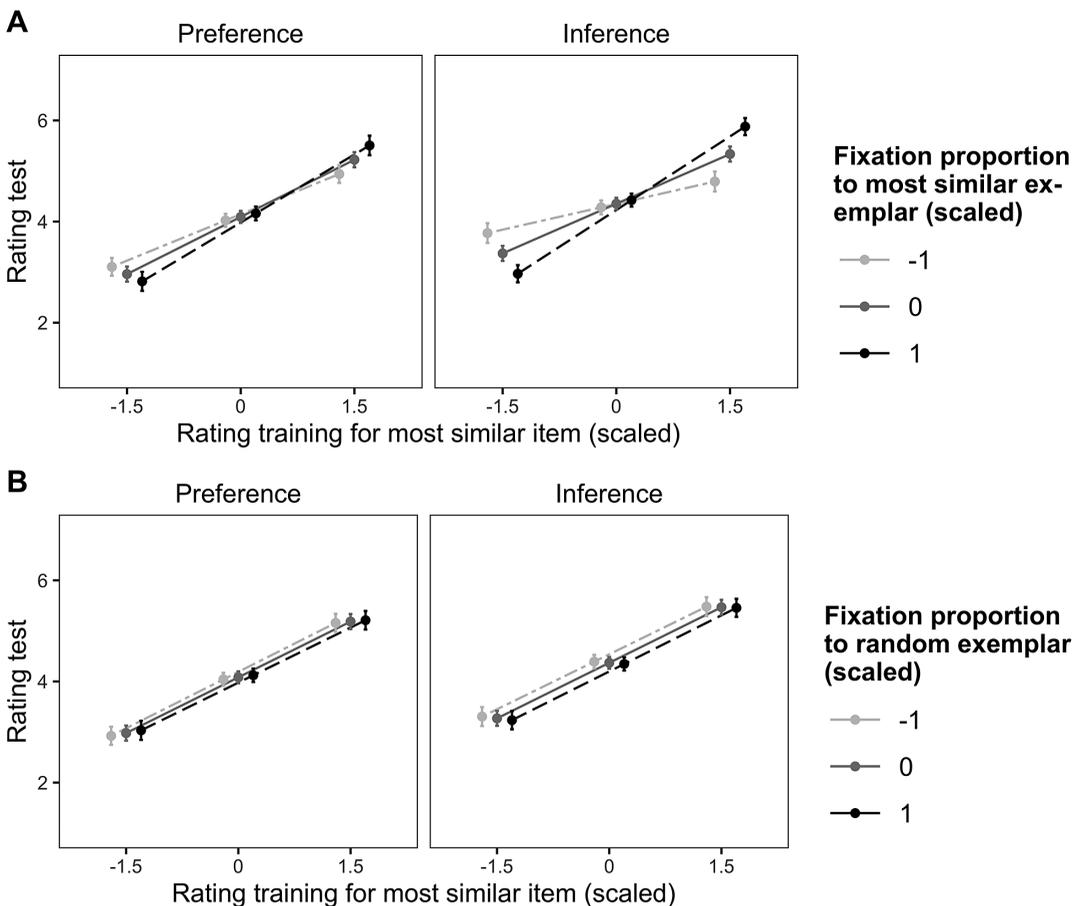


Figure 4. Estimated marginal means for the influence of training rating of the most similar exemplar on the test rating. This influence is moderated by the amount of looking at nothing (A). When randomly selecting fixation proportions for any of the exemplar locations, the moderating effect goes away (B). The variables training rating and fixation proportions were centered on 0. A value of ± 1 indicates ± 1 SD. Standard errors show estimated 95% confidence intervals.

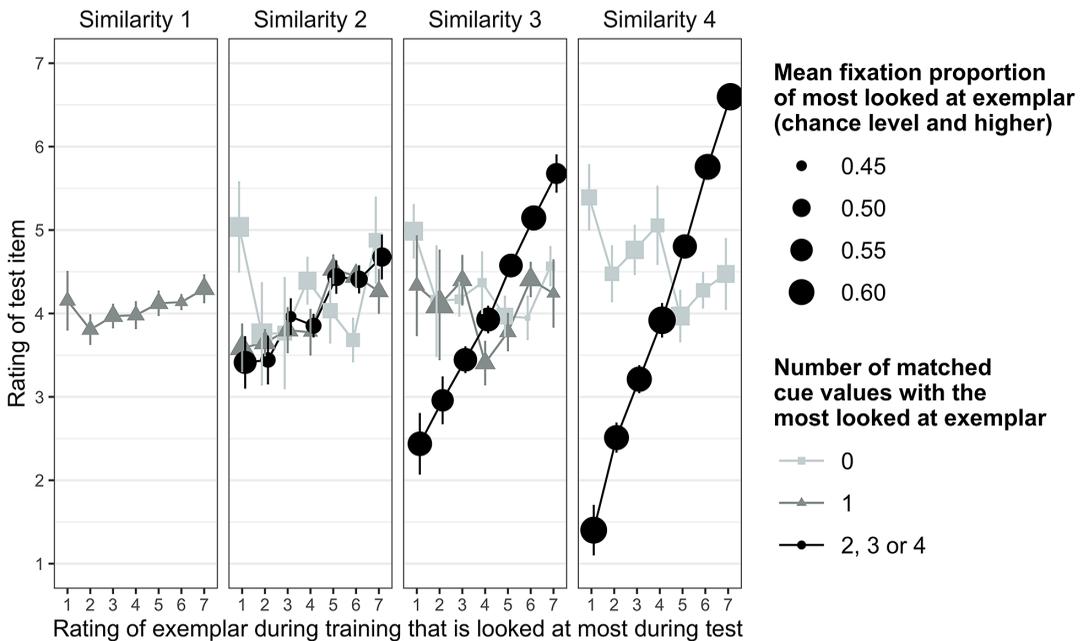


Figure 5. Mean test ratings plotted by training rating of the exemplar that is looked at most (x-axis), similarity (panels 1 to 4 from left to right), and number of matches in cue values between the most looked at exemplar and the test item (e.g., Panel 2 summarizes data for items that shared 2 cue values with 1 of the 4 exemplars and 0 or 1 cue values with each of the other 3 exemplars). Darker points indicate more matches in cue values. The size of points indicates the strength of looking at nothing with larger points indicating higher fixation proportions. Results are plotted over task conditions. Standard errors show within-subject 95% confidence intervals.

In sum, the more participants looked at the location of the most similar training exemplar, the more their ratings resembled the rating they gave the most similar exemplar during training. This is the case for both, the inferential and the preferential judgment condition, even though the effect is more pronounced in the inferential condition.

Thus, when the most similar exemplar is looked at, looking more increases the link between the exemplar's criterion value and the judgment. However, a more direct test of the relationship between looking at nothing and judgments is to analyze whether the training rating of the exemplar they look at most during each test trial predicts the test rating of this trial, which we tested in the second analysis.

2.2.4.2. Does looking the most at an exemplar predict judgments?

Figure 5 shows that when the most looked at training exemplar is highly similar to the test item, test judgments are closely related to the training rating it received (panels 3 and 4). However, in trials, in which the most looked at training item has only a low number of matches in cue values with the test item, the correlation vanishes (panel 1). To test this pattern, we ran a linear mixed model for test ratings with random intercepts for subjects and exemplars and the predictors task condition and similarity. We added 2 new variables and their interactions to this analysis. For one, the rating of the exemplar that was looked at most during the test (ranging from 1 to 7). Secondly, we added the number of matched cue values between the exemplar looked at most and the test item (ranging from 0 to 4 matches). For instance, in a trial with a test item that shared 3 cue values with one training exemplar, a participant could have either looked at the exemplar with 3 matches in cue values (i.e., the most similar exemplar location), the exemplar that shared only 1 cue value with the test item, or at 1 of the 2 remaining

exemplars sharing 0 cue values with the test item. To determine, which exemplar was looked at most, in the analysis, we only included trials in which participants looked more than chance level to any 1 of the 4 exemplar locations (fixation proportions to most looked-at location larger than 0.25). To determine, which exemplar location received the most attention during the judgment process, the calculation of fixation proportions also contained fixation proportions outside the 4 exemplar AOIs. Results show a main effect of the rating of the most looked-at exemplar on the test ratings: $F(1, 2165.19) = 320.72$, $p < .001$. The analyses also revealed a significant interaction between the rating of the most looked-at exemplar and the number of matches in cue values between the most looked-at exemplar and the test item, $F(4, 2345.53) = 88.94$, $p < .001$. Figure 5 shows that in cases when people look the most at the exemplar location of an exemplar with at least 1 match up to 4 matches in cue values, eye movements predict judgments (linear contrasts all $ps < .001$). However, when looking the most at an exemplar with 0 matches in cue values there is no relation between the rating of that exemplar and the training rating (linear contrast $p = .68$). There was no significant effect of task condition, $F(1, 74.30) = 3.79$, $p = .055$. No other main effect or interaction reached significance (all $ps > .084$).

Together, these analyses show that looking-at-nothing to exemplar locations predicts ratings, but only when people looked at the most similar exemplar location. When looking the most at the most dissimilar exemplar location, the training rating of the most looked-at but dissimilar exemplar does not predict the test rating.

2.3. Discussion

The goal of this study was to test whether exemplar retrieval also occurs during preferential judgments and to link eye movements to preferential judgments. Eye-movement measures based on the looking-at-nothing behavior have been shown to indicate exemplar retrieval from memory during inferential choice (Rosner et al., 2022; Rosner & von Helversen, 2019; Scholz et al., 2015). In line with previous research, we found that the amount people looked at empty spatial locations reflected the similarity (defined as a number of matches in cue values) between test items and training exemplars. Importantly, we found the same pattern for preferential judgments, where there were no correct or wrong response options and the evaluations depended entirely on the individual preferences of the participants. This suggests that also in preferential judgments, people retrieve information about previously seen exemplars from memory. It must be noted, however, that the influence of similarity on looking-at-nothing behavior was reduced in comparison to inferential judgments, which could indicate a reduced role of exemplar-based processes in preferential judgments. One potential limitation of our experiment is that we instructed participants in the preferential condition to respond as consistently as possible. Nevertheless, individual differences in consistency could not eliminate the effects of the task condition and similarity manipulations on eye movements. It may still be the case that participants in the preference condition retrieved exemplars from memory only because we instructed them in the test phase to respond as consistently as possible with their judgments during the training phase. Thus, it is important to show evidence for retrieval even when participants are not instructed to give consistent judgments.

Another limitation may arise from the fact that we assigned the first 40 participants to the preference condition and the second 40 participants to the inference condition. As subjects in the inference condition were presented with ratings from subjects of the preference condition, the preference condition had to be assessed first. However, all participants taking part in the experiment stem from the same population and the time point at which participants signed up for the experiment seemed random. Thus, we do not anticipate any systematic sampling bias in this experiment.

Regarding the role of exemplar processes during judgment, we found that eye movements to the most similar exemplar moderated the influence of the training rating on the test rating, suggesting a closer link the longer an exemplar was fixated. However, the analyses on most looked-at exemplar locations revealed a pattern that cannot be explained by a strategy assuming that judgments are purely based on the retrieval of past instances stored in memory. In trials, in which participants mostly looked

at training exemplars with a low similarity, we found no influence of the rating of this exemplar on the rating of the test item. If eye movements indicate retrieval and retrieval is directly translated into a judgment, one would expect a positive correlation even when similarity is low. Thus, these results suggest that participants were relying on further judgment processes besides exemplar retrieval. We will return to this question in the general discussion.

As an alternative process to judgment, people may not have relied on the retrieval of exemplars. Instead, they could have used an ingredient-based strategy. An ingredient-based strategy may also involve memory retrieval, albeit most likely less than when making judgments based on exemplar similarity. For instance, in the decision-by-sampling approach (Stewart et al., 2006) the evaluation of attributes such as ingredients is based on the comparison of attribute values sampled from memory. Thus, eye movements to the locations may not reflect retrieval of the exemplar, but retrieval processes related to the evaluation of single ingredients.

We conducted a second experiment to ensure that our results are replicable and to test alternative explanations such as memory processes due to the evaluation of ingredients.

3. Experiment 2

In Experiment 2, we studied looking-at-nothing only in preferential judgments. To rule out the alternative explanation that participants showed the observed eye-movement effects only because they were instructed to respond as consistently as possible (i.e., according to their judgments in the training phase), half the participants received the same instructions as in Experiment 1 (preference consistency condition) and the other half received no consistency instructions (preference condition). Additionally, we investigated whether the eye movements could also be explained by a judgment strategy that integrates the values of single ingredients. For this, we adapted the ingredient-rating phase by presenting ingredients auditorily, leaving the screen blank, and recording eye movements. If people show looking-at-nothing behavior when rating their preferences for single ingredients, this might explain the looking-at-nothing behavior while evaluating the test, which would suggest that mechanisms other than exemplar retrieval underlie the preferential judgments. All experimental materials, data, analysis scripts, can be found on OSF (<https://osf.io/fjmkd/>).

3.1. Method

3.1.1. Participants

Sixty-seven participants (50 female, $M_{\text{age}} = 25.5$ years, range 18–55 years) at the University of Zurich took part in the study for course credit or financial compensation (15 CHF per hour). All participants had normal or corrected-to-normal vision. Mean tracking accuracy in the test trials was very high at $M = 0.7^\circ$ of visual angle. All participants signed informed consent forms. The preference consistency condition had 36 participants and the preference condition had 31.

3.1.2. Materials and apparatus

The same materials and apparatus as in Experiment 1 were used.

3.1.3. Procedure

The procedure in both conditions of Experiment 2 followed that of the preference condition of Experiment 1. That is, after an initial test of the eye tracker and practice with the rating scale, participants were familiarized with the experimental materials during the training phase. Note that in Experiment 2, we showed only 3 repetitions per exemplar. In the test phase, participants had to judge the 40 smoothies (Appendix A). Therefore, participants in the preference consistency condition received the same instructions as participants in the preference condition of Experiment 1. Participants in the preference condition received no consistency instructions. After the test phase, we assessed location memory for ingredients. The procedure of the ingredient-rating phase differed from that of

Table 3. Results of behavioral measures in Experiment 2.

Variable	Condition				<i>b</i>	<i>t</i> (65)	<i>p</i>	<i>R</i> ²
	Preference consistency		Preference					
	<i>M</i> (<i>SD</i>)	95% CI	<i>M</i> (<i>SD</i>)	95% CI				
Preference judgments								
Exemplars	4.43 (0.58)	[4.2, 4.6]	4.83 (0.70)	[4.6, 5.1]	0.40	2.56	.013	.08
Ingredients	4.61 (0.42)	[4.4, 4.8]	5.00 (0.55)	[4.8, 5.2]	0.40	3.32	.002	.13
Learn–test consistency	0.53 (0.34)	[0.4, 0.6]	0.57 (0.31)	[0.5, 0.7]	0.04	0.50	.619	-.01
Location memory test								
Exemplar level	11.5 (3.41)	[10.4, 12.6]	11.5 (3.03)	[10.3, 12.6]	–0.08	–0.10	.924	–0.02
Ingredients level	7.11 (4.12)	[5.8, 8.4]	6.39 (3.86)	[5.0, 7.8]	–0.72	–0.74	.46	–0.01

Note: CI, confidence interval.

Experiment 1. It closely followed the test phase procedure of both experiments (Figure 2). That is, participants saw the empty rectangles that had contained the ingredients on the screen. One ingredient at a time was read out loud. When participants were ready to enter their judgment, they could proceed to the rating scale with a mouse click. At the end of the experiment, participants worked through the postquestionnaire asking them about food allergies and their smoothie consumption habits. Experiment 2 lasted on average 35 min.

3.2. Results

First, we provide an overview of the behavioral measures, before reporting results on the similarity effect and the link between judgments and eye movements. The analyses followed the same rationale as in Experiment 1.

3.2.1. Behavioral measures

The postquestionnaire revealed no serious food allergies concerning the ingredients used in the study. All participants⁶ liked smoothies (all values larger than 1) and a large part of the sample regularly drank smoothies at least once per month (0 times: *n* = 4, 1–3 times: *n* = 31, 4–6 times: *n* = 17, more than 6 times: *n* = 7).

Also, in Experiment 2, preferences for exemplars and ingredients ranged over the whole rating scale (range = 1–7). Appendix C contains an overview of the ratings on the level of single exemplars and ingredients. For both preference measures, participants in the preference condition had an overall slightly higher rating of how much they liked the presented exemplars in comparison to participants in the preference consistency condition (Table 3). As we held no hypothesis on how liking could affect the looking-at-nothing behavior, we did not include preference ratings as covariates in the analyses of eye-movement behavior.

Participants in the 2 conditions had similar results for the task concerning learn–test consistency and location memory performance (Table 3) and performed similarly to participants in the preference condition of Experiment 1.⁷ Although we did not observe differences in conditions, for reasons of comparability with the analyses of Experiment 1, we also added them as covariates to the analyses of eye movements in Experiment 2.

⁶Note that only 59 out of 67 participants responded to the questions.

⁷Results on test-test consistency of Experiment 2 can be found in the Supplementary Material. When controlling for the guessing rate (Budescu & Bar-Hillel, 1993), participants still performed better when analyzing location memory performance on the level of exemplars than the level of ingredients.

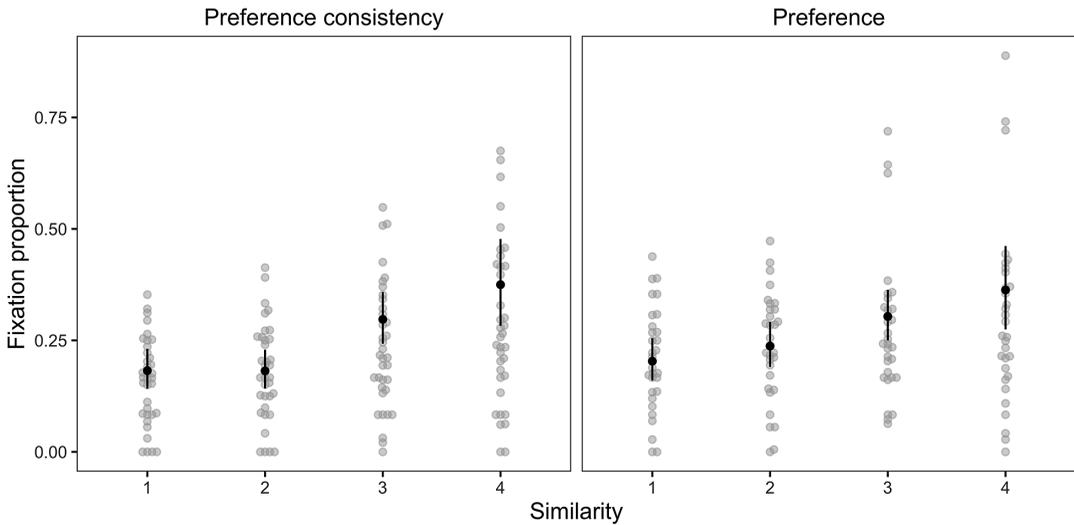


Figure 6. Estimated means for proportions of fixation on the most similar exemplar location as a function of the similarity of test items to exemplars for the 2 task conditions in Experiment 2. Standard errors show estimated within-subject 95% confidence intervals. Gray dots show individual participants' means.

3.2.2. Similarity effect on eye movements

As in Experiment 1, we analyzed how the similarity between all test items and training exemplars affected looking-at-nothing behavior. Furthermore, we tested for differences between the instruction conditions (with and without instructions to judge as consistently as possible during preferential choices). We added instruction condition and similarity as well as their interaction as fixed effects for the prediction of proportions of fixation on the most similar exemplar location to a GLMM for binomially distributed data. The random effects structure consisted of by-subject random intercepts and by-subject random slopes for exemplars and similarity. We ran 2 models that either added learn-test consistency or location memory performance as covariates. Concerning the first model, we found a main effect of similarity, $\chi^2(3) = 27.67, p < .001$. There was no main effect of task condition, $\chi^2(1) = 0.55, p = .46$, nor an interaction between task condition and similarity, $\chi^2(3) = 2.88, p = .41$ (Figure 6), or an influence of the learn-test consistency. Pairwise comparisons with Holm correction showed significant differences between all levels of the factor similarity (all $ps < .047$), except the difference between similarity levels 1 and 2 ($p = .512$).⁸

Adding number of correct retrievals in the location memory test revealed the same result pattern and a significant effect on location memory performance, $\chi^2(1) = 6.78, p = .009$. That is, the better participants' location memories, the higher the proportion of fixations on the relevant exemplar location (linear contrast: $z = 2.64, p = .008$).

3.2.3. Eye movements and judgments

3.2.3.1. Looking at nothing moderates judgments

We repeated the same analyses on the moderating influence of eye movements on the relation between training and test ratings with the 2 preference conditions. Like in Experiment 1, we find a significant influence of the training rating on the test rating (Table 4). The more participants looked at the exemplar locations, the more the training rating of the most similar exemplar resembled the test rating (Figure 7A). We found no differences between task conditions.

⁸Running the models on the similarity effect on eye movements without covariates revealed the same results pattern. The same is true when running the model without items being identical to a training item (similarity = 4).

Table 4. Results of the mixed model analyses test the moderating effect of looking at nothing on the relation between training and test judgments of Experiment 2.

Fixed effect	<i>b</i>	<i>SE</i>	<i>F</i>	<i>df</i>	<i>p</i>
(Intercept)	4.00	0.10			
Training rating	0.68	0.03	542.91	1, 2768.26	<.001
Fixation proportions	0.11	0.03	15.10	1, 3199.83	<.001
Task condition	-0.08	0.06	2.13	1, 65.44	.149
Training rating × Fixation proportions	0.11	0.03	15.59	1, 3196.04	<.001
Training rating × Task condition	-0.01	0.03	0.03	1, 3163.54	.960
Fixation proportions × Task condition	0.04	0.03	1.66	1, 3199.92	.198
Training rating × Fixation proportions × Task condition	-0.01	0.03	0.20	1, 3195.37	.657

Note: CI, confidence interval. Task conditions were coded as 1 = Preference consistency, 2 = Preference.

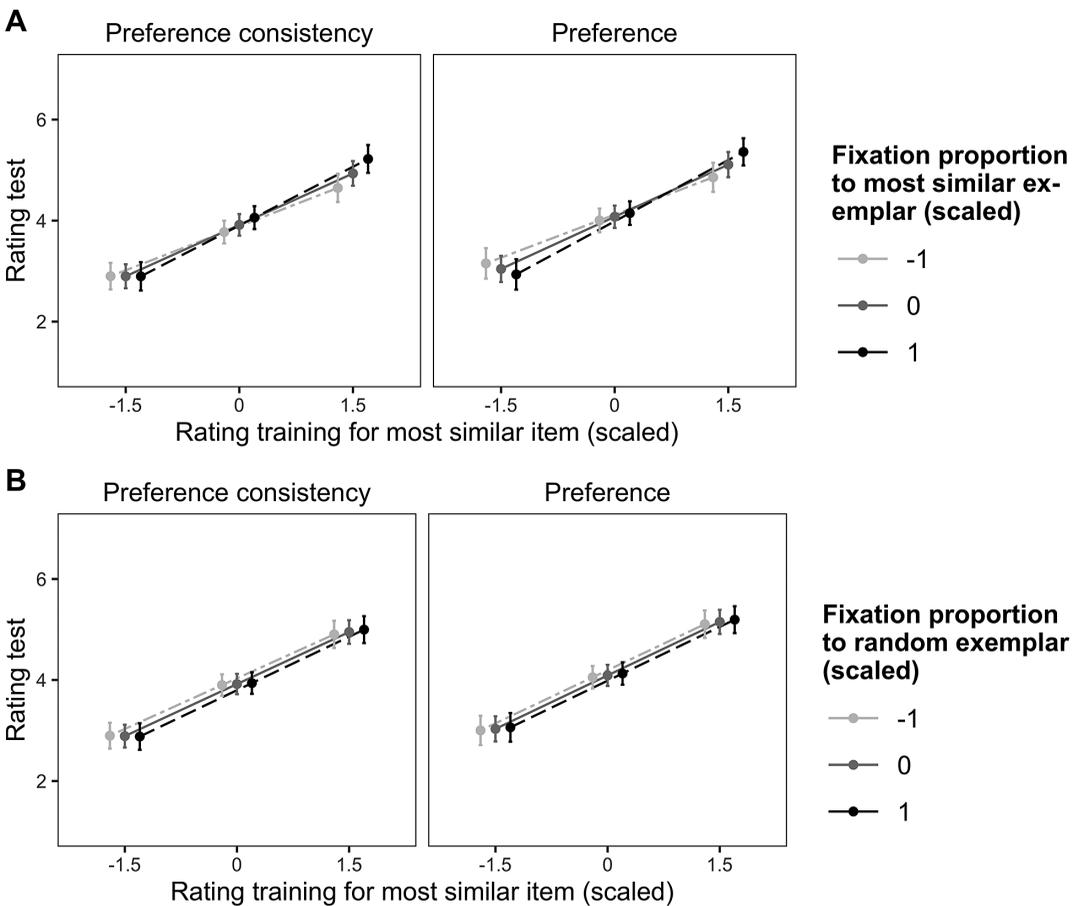


Figure 7. Estimated marginal means for the influence of training rating of the most similar exemplar on the test rating. This influence is moderated by the amount of looking at nothing (A). When randomly selecting the fixation proportions for any of the exemplar locations, the moderating effect goes away (B). The variables training rating and fixation proportions were centered on 0. Standard errors show estimated 95% confidence intervals.

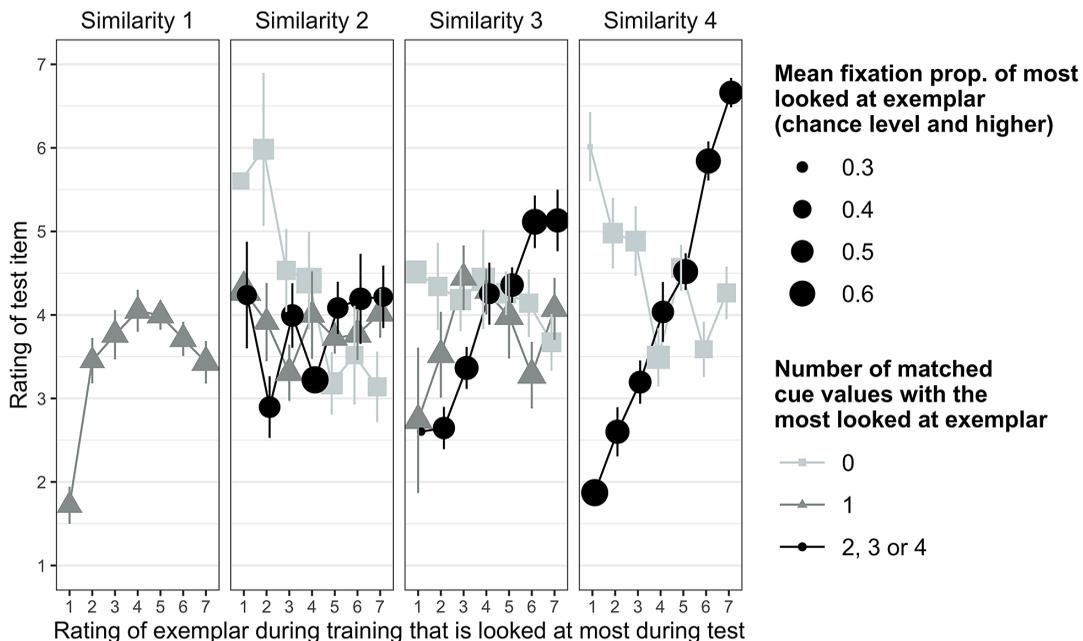


Figure 8. Mean test ratings are plotted by training rating of the exemplar that is looked at most (*x*-axis), similarity, and number of matches in cue values between the most looked at exemplar and the test item. Darker points indicate more matches in cue values. The size of points indicates the strength of looking at nothing with larger points indicating higher fixation proportions. Results are plotted over task condition. Standard errors show within-subject 95% confidence intervals.

Like in Experiment 1, when adding fixation proportions to a randomly chosen exemplar as a predictor, the moderating influence goes away, $F(1, 3172.15) = 0.20, p = .651$ (Figure 7B).

3.2.3.2. Does looking the most at an exemplar predict judgments?

Like in Experiment 1, we tested, if the learning rating of the exemplar looked at most during each trial of the test phase predicts the resulting test rating. We conducted the same analyses as for Experiment 1, but for the 2 preference conditions of Experiment 2. Overall, the pattern of results is the same (Figure 8). We find a main effect for the rating of the exemplar looked at most predicting test ratings, $F(1, 1215.73) = 56.32, p < .001$. And also a significant interaction between the rating of the most looked at exemplar and the number of matches in cue values between the most looked at exemplar and the test item, $F(4, 1249.89) = 34.47, p < .001$. When looking the most to highly similar exemplars, the exemplar looked at most predicted the judgment (linear contrasts for 2 to 4 matches $ps < .05$). When looking at less similar exemplars, either no relation between test and training judgments (linear contrast 1 match $p = .15$) or even a negative relation was observed (linear contrast 0 matches $p = .001$). The 2 instruction conditions did not differ, $F(1, 62.32) = 3.34, p = .072$.

3.2.3.3. No looking-at-nothing during ingredients rating

Last, we tested, if looking at nothing also occurs when rating preferences for individual ingredients during the ingredient rating phase at the end of this experiment. Participants exhibited almost no looking at nothing during ingredients ratings (fixation proportions to the exemplar location that had contained the tested ingredient: $M_{\text{Pref cons.}} = 0.08, SD_{\text{Pref cons.}} = 0.08, M_{\text{Pref}} = 0.06, SD_{\text{Pref}} = 0.06$) and that did not differ between the instruction conditions, $\chi^2(1) = 2.33, p = .127$. That is, although people were able to retrieve the locations of the ingredients when evaluating how much they like each ingredient, there is no evidence that they retrieved this information.

3.3. Discussion

The goal of the second experiment was to rule out the alternative explanation that the similarity effect in a preferential judgment task occurred only because participants were instructed to respond as consistently as possible. We found that people still showed looking-at-nothing behavior during preferential judgments when there were no instructions to respond as consistently as possible.

Additionally, we explored if participants looked at nothing when judging their preferences for single ingredients. We could not find looking-at-nothing during the ingredients rating phase. Thus, evaluating how much people like an ingredient does not draw on memory retrieval to the same extent as evaluating how much one would like smoothies varying in their similarity to previously encountered exemplars stored in memory.

4. General discussion

This study used eye movements to learn about exemplar retrieval during inferential and preferential judgment. In preferential judgments, people rate an object based on its subjective utility. Thus, there is no objective criterion the judgment can be compared with. Here, eye movements can be a useful tool to gain process-level evidence for the processes underlying the judgment.

Experiment 1 compared exemplar retrieval between an inferential and a preferential judgment task. Looking-at-nothing reflected the similarity (defined as the number of matches in cue values) between test items and training exemplars, that is, the similarity effect on eye movements. The similarity effect on eye movements also occurred during preferential judgments. Experiment 2 replicated the finding for preferential judgments and ruled out the alternative hypothesis that the observed eye-movement effects occurred only because participants were instructed to respond as consistently as possible. These findings are in line with previous research showing that as the similarity between training and test items increases, people proportionally look more to the associated but emptied spatial locations (Rosner et al., 2022; Rosner & von Helversen, 2019; Scholz et al., 2015). However, does this similarity effect on eye movements mean that people retrieve information about exemplars from memory? In favor of the exemplar retrieval hypothesis speaks that looking-at-nothing behavior has been strongly connected to the retrieval of information from memory in a variety of memory (Johansson et al., 2006; Martarelli & Mast, 2013; Richardson & Spivey, 2000; Scholz et al., 2016) and decision-making tasks (Pärnamets et al., 2016; Platzer et al., 2014; Renkewitz & Jahn, 2012; Rosner et al., 2022; Rosner & von Helversen, 2019). However, one might argue that eye movements may rather reflect the retrieval of single attribute information (i.e., ingredients). Against this explanation speaks that Scholz et al. (2015) showed in a comparable design that relying on an exemplar-based strategy induced the observed eye movement behavior, while applying a simple rule did not. In addition, we found that the retrieval of the ratings of how much one would like each single ingredient at the end of Experiment 2 did not elicit looking-at-nothing behavior—where according to this explanation one would expect a similar amount—while at the same time, people were still very accurate in their location memories for where ingredients have been presented during training. Last, location memory performance was higher when analyzing it on the level of exemplars than on the level of single ingredients hinting toward information rather being represented on the level of exemplars. Nevertheless, future research is needed to test to what extent looking-at-nothing reflects the retrieval of exemplars from memory.

If we assume that eye movements reflect the retrieval of information about the looked-at exemplars, however, it raises the question of how the pattern of results connecting eye movements with judgments can be explained. Here, we found for both, the inferential and preferential judgments, firstly that ratings of exemplars during training predicted test ratings and that this relation was moderated by the looking-at-nothing behavior. The more strongly people looked at exemplars during test, the more their test ratings resembled the training rating of the respective exemplar. Secondly, we found that training ratings of the exemplar looked at most predicted test ratings when people looked at exemplars that were highly similar to the test item, but not when people looked at dissimilar exemplars. The first result is in line with Rosner and von Helversen (2019) who found that a biasing effect of retrieving exemplars

was mediated by the amount of looking at the exemplar locations. Furthermore, when adding randomly selected fixation proportions as a moderator to the analysis, the moderating effect diminishes. This finding supports the idea that eye movements reflect exemplar retrieval and retrieved exemplars are used to form judgments. However, our second results speak against the idea that exemplar retrieval is the sole determinant of judgments, given that the rating of the most looked at exemplar only predicted test judgments when participants looked at a highly similar exemplar. This result is, however, in line with previous accounts on the use of exemplar- and rule-processes in judgment, assuming that judgments are the result of a mixture of processes or strategy blending (Albrecht et al., 2020; Herzog & von Helversen, 2018; Izydorczyk & Bröder, 2021). For instance, if participants retrieve a training exemplar, then consider the difference between the retrieved exemplar and the test item and adjust their judgment accordingly as suggested by Albrecht et al. (2020) it could explain why we find evidence of retrieval on the one hand but only a correlation between the rating of the retrieved exemplar and the test rating when the similarity between test item and retrieved exemplar is high. In this study design, it is difficult to further tease apart the processes that may have produced the observed judgment. Furthermore, this interpretation is based on exploratory analyses. Additional experiments may be necessary to replicate the observed findings. Furthermore, future research might seek to design experiments that involve testing eye movements combined with modeling judgment strategies, despite the inherent difficulty in the domain of preferential judgments (but see Jarecki & Rieskamp, 2022 for a fruitful approach).

Taken together, the results provide behavioral evidence that exemplars are retrieved from memory even in preferential judgments. However, it must be noted that the effect was reduced in comparison to when making an inferential judgment. This result suggests that the retrieval of exemplars did not occur to the same extent as during inferential judgments. There are several reasons that could underlie this finding. For instance, it is possible that the similarity effect was reduced for preferential choices because in some cases, participants may have had immediate emotional responses to ingredients that determined their preferences (Loewenstein et al., 2001; Zajonc, 1980, 2000) and thus may have had less need to retrieve complete memory traces (Betsch et al., 2001). Alternatively, it is possible that people rely to the same degree (or even more) on exemplars when making judgments but just not on the exemplars presented to them during the study. In the inference condition, participants had no further information beyond the presented exemplars to infer the other person's preference and thus it is likely that they retrieved the exemplars to make their judgment. However, when providing their own preferences people may also have retrieved exemplars experienced outside of the study (e.g., following habitual behavior, Verplanken & Orbell, 2022). Thus, even though the presented exemplars were recent and relevant, they may have been retrieved less due to competing previous memory traces. One way to disentangle these explanations—though not that easy to achieve—would be to use stimulus material with which people have had no previous experience.

None of the behavioral measures that we analyzed as covariates could alter or alternatively explain our findings. Still, they provide useful insights about the looking-at-nothing behavior during multiple-cue judgments. Looking-at-nothing was more pronounced when people had better memories about the exemplar locations, and in Experiment 1 when they responded more consistently. Previous research has argued that better memory representations are related to more exemplar-based decision-making (Hoffmann et al., 2013). Thus, high memory activation of exemplar information may go along with better retrieval accuracy and more looking-at-nothing. Consequently, if the information is more readily available, it may also be easier to give more consistent responses.

The current study aimed to design a situation in which exemplar retrieval is a likely mechanism involved in solving the task and to use eye movements to trace the assumed retrieval processes. On the one hand, such a design is warranted to carefully draw conclusions about the assumed memory mechanisms (Schoemann et al., 2019). This may, however, limit the ecological validity of the observed behavioral and eye movement patterns. Thus, future research will have to test to what extent these results can be generalized to other situations and contexts.

In sum, we found that also in preferential judgment tasks, people showed eye movements to the locations of previously presented exemplars. In inference tasks these eye movements have been

linked to exemplar retrieval, suggesting that people may also rely on exemplar retrieval when making preferential judgments. This result highlights (1) that inferential and preferential judgments seem to have many commonalities concerning the underlying retrieval processes and (2) the usefulness of studying eye movements as a process measure to get direct behavioral evidence of the memory processes of interest.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/jdm.2024.3>.

Data availability statement. All materials and data are available at: <https://osf.io/fjmkd/>.

Acknowledgments. All authors gratefully acknowledge the support of the Swiss National Science Foundation (SNSF grant 157432). The first author received additional funding from SNSF grant 186032. Furthermore, the authors thank Michael Schaffner for programming Experiment 2 and together with Michelle Sedlak for collecting the data of Experiment 2.

References

- Aarts, H., Verplanken, B., & Van Knippenberg, A. (1998). Predicting behavior from actions in the past: Repeated decision making or a matter of habit? *Journal of Applied Social Psychology, 28*(15), 1355–1374. <https://doi.org/10.1111/j.1559-1816.1998.tb01681.x>
- Albrecht, R., Hoffmann, J. A., Pleskac, T. J., Rieskamp, J., & von Helversen, B. (2020). Competitive retrieval strategy causes multimodal response distributions in multiple-cue judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(6), 1064–1090. <https://doi.org/10.1037/xlm0000772>
- Altmann, G. T. M. (2004). Language-mediated eye movements in the absence of a visual world: The ‘blank screen paradigm.’ *Cognition, 93*(2), B79–B87. <https://doi.org/10.1016/j.cognition.2004.02.005>
- Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences, 16*(8), 437–443. <https://doi.org/10.1016/j.tics.2012.06.010>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Basu, K. (1993). Consumers’ categorization processes: An examination with two alternative methodological paradigms. *Journal of Consumer Psychology, 2*(2), 97–121. [https://doi.org/10.1016/S1057-7408\(08\)80020-4](https://doi.org/10.1016/S1057-7408(08)80020-4)
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Betsch, T., Plessner, H., Schwieren, C., & Gütig, R. (2001). I like it but I don’t know why: A value-account approach to implicit attitude formation. *Personality and Social Psychology Bulletin, 27*(2), 242–253. <https://doi.org/10.1177/0146167201272009>
- Bordalo, P., Gennaioli, N., & Shleifer, A. (2020). Memory, attention, and choice. *The Quarterly Journal of Economics, 135*(3), 1399–1442. <https://doi.org/10.1093/qje/qjaa007>
- Brehmer, B. (1994). The psychology of linear judgement models. *Acta Psychologica, 87*(2–3), 137–154. [https://doi.org/10.1016/0001-6918\(94\)90048-5](https://doi.org/10.1016/0001-6918(94)90048-5)
- Bröder, A., & Gräf, M. (2018). Retrieval from memory and cue complexity both trigger exemplar-based processes in judgment. *Journal of Cognitive Psychology, 30*(4), 406–417. <https://doi.org/10.1080/20445911.2018.1444613>
- Bröder, A., Gräf, M., & Kieslich, P. J. (2017). Measuring the relative contributions of rule-based and exemplar-based processes in judgment: Validation of a simple model. *Judgment and Decision Making, 12*(5), 491. <https://doi.org/10.1017/s1930297500006513>
- Bröder, A., Newell, B. R., & Platzer, C. (2010). Cue integration vs. exemplar-based reasoning in multi-attribute decisions from memory: A matter of cue representation. *Judgment and Decision Making, 5*(5), 326–338. <https://doi.org/10.1017/S1930297500002138>
- Budescu, D., & Bar-Hillel, M. (1993). To guess or not to guess: A decision-theoretic view of formula scoring. *Journal of Educational Measurement, 30*(4), 277–291.
- Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127*(2), 107–140. <https://doi.org/10.1037/0096-3445.127.2.107>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Ferreira, F., Apel, J., & Henderson, J. M. (2008). Taking a new look at looking at nothing. *Trends in Cognitive Sciences, 12*(11), 405–410. <https://doi.org/10.1016/j.tics.2008.07.007>
- Gilboa, I., & Schmeidler, D. (2001). A theory of case-based decisions. In *A theory of case-based decisions*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511493539>
- Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science, 27*(4), 591–635. [https://doi.org/10.1016/S0364-0213\(03\)00031-4](https://doi.org/10.1016/S0364-0213(03)00031-4)
- Grosskopf, B., Sarin, R., & Watson, E. (2015). An experiment on case-based decision making. *Theory and Decision, 79*(4), 639–666. <https://doi.org/10.1007/s11238-015-9492-1>

- Hahn, U., Prat-Sala, M., Pothos, E. M., & Brumby, D. P. (2010). Exemplar similarity and rule application. *Cognition*, 114(1), 1–18. <https://doi.org/10.1016/j.cognition.2009.08.011>
- Hedge, C., Oberauer, K., & Leonards, U. (2015). Selection in spatial working memory is independent of perceptual selective attention, but they interact in a shared spatial priority map. *Attention, Perception, and Psychophysics*, 77(8), 2653–2668. <https://doi.org/10.3758/s13414-015-0976-4>
- Herzog, S. M., & von Helversen, B. (2018). Strategy selection versus strategy blending: A predictive perspective on single- and multi-strategy accounts in multiple-cue estimation. *Journal of Behavioral Decision Making*, 31(2), 233–249. <https://doi.org/10.1002/bdm.1958>
- Hoffmann, J. A., von Helversen, B., & Rieskamp, J. (2013). Deliberation's blindsight: How cognitive load can improve judgments. *Psychological Science*, 24(6), 869–879. <https://doi.org/10.1177/0956797612463581>
- Hoffmann, J. A., von Helversen, B., & Rieskamp, J. (2014). Pillars of judgment: How memory abilities affect performance in rule-based and exemplar-based judgments. *Journal of Experimental Psychology: General*, 143(6), 2242–2261. <https://doi.org/10.1037/a0037989>
- Hoffmann, J. A., von Helversen, B., & Rieskamp, J. (2016). Similar task features shape judgment and categorization processes. *Journal of Experimental Psychology: Learning Memory and Cognition*, 42(8), 1193–1217. <https://doi.org/10.1037/xlm0000241>
- Huettig, F., Mishra, R. K., & Olivers, C. N. L. (2012). Mechanisms and representations of language-mediated visual attention. *Frontiers in Psychology*, 2, 1–11. <https://doi.org/10.3389/fpsyg.2011.00394>
- Izidorczyk, D., & Bröder, A. (2021). Exemplar-based judgment or direct recall: On a problematic procedure for estimating parameters in exemplar models of quantitative judgment. *Psychonomic Bulletin & Review*, 28(5), 1495–1513. <https://doi.org/10.3758/s13423-020-01861-1>
- Jahn, G., & Braatz, J. (2014). Memory indexing of sequential symptom processing in diagnostic reasoning. *Cognitive Psychology*, 68, 59–97. <https://doi.org/10.1016/j.cogpsych.2013.11.002>
- Jarecki, J. B., & Rieskamp, J. (2022). Comparing attribute-based and memory-based preferential choice. *Decision*, 49, 65–90. <https://doi.org/10.1007/s40622-021-00302-9>
- Johansson, R., Holsanova, J., & Holmqvist, K. (2006). Pictures and spoken descriptions elicit similar eye movements during mental imagery, both in light and in complete darkness. *Cognitive Science*, 30(6), 1053–1079. https://doi.org/10.1207/s15516709cog0000_86
- Juslin, P., Karlsson, L., & Olsson, H. (2008). Information integration in multiple cue judgment: A division of labor hypothesis. *Cognition*, 106(1), 259–298. <https://doi.org/10.1016/j.cognition.2007.02.003>
- Juslin, P., Olsson, H., & Olsson, A. C. (2003). Exemplar effects in categorization and multiple-cue judgment. *Journal of Experimental Psychology: General*, 132(1), 133–156. <https://doi.org/10.1037/0096-3445.132.1.133>
- Karlsson, L., Juslin, P., & Olsson, H. (2008). Exemplar-based inference in multi-attribute decision making: Contingent, not automatic, strategy shifts? *Judgment and Decision Making*, 3(3), 244–260. <https://doi.org/10.1017/S1930297500002448>
- Keeney, R. L., Raiffa, H., & Rajala, D. W. (1979). Decisions with multiple objectives: Preferences and value trade-offs. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(7), 403–403. <https://doi.org/10.1109/TSMC.1979.4310245>
- Klichowicz, A., Lippoldt, D. E., Rosner, A., & Krems, J. F. (2021). Information stored in memory affects abductive reasoning. *Psychological Research*, 85(8), 3119–3133. <https://doi.org/10.1007/s00426-020-01460-8>
- Krefeld-Schwab, A., & Rosner, A. (2020). A new way to guide consumer's choice: Retro-cueing alters the availability of product information in memory. *Journal of Business Research*, 111, 135–147. <https://doi.org/10.1016/j.jbusres.2019.08.012>
- Laeng, B., Bloem, I. M., D'Ascenzo, S., & Tommasi, L. (2014). Scrutinizing visual images: The role of gaze in mental imagery and memory. *Cognition*, 131(2), 263–283. <https://doi.org/10.1016/j.cognition.2014.01.003>
- Lajos, J., Katona, Z., Chattopadhyay, A., & Sarvary, M. (2009). Category activation model: A spreading activation network model of subcategory positioning when categorization uncertainty is high. *Journal of Consumer Research*, 36(1), 122–136. <https://doi.org/10.1086/595024>
- Lenth, R. V. (2022). *emmeans: Estimated marginal means, aka least-squares means* (R package version 1.7.5.). <https://cran.r-project.org/package=emmeans>
- Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, 127(2), 267–286. <https://doi.org/10.1037/0033-2909.127.2.267>
- Martarelli, C. S., & Mast, F. W. (2013). Eye movements during long-term pictorial recall. *Psychological Research*, 77(3), 303–309. <https://doi.org/10.1007/s00426-012-0439-7>
- Martindale, C. (1984). The pleasures of thought: A theory of cognitive hedonics. *Journal of Mind and Behavior*, 5(1), 49–80.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85(3), 207–238. <https://doi.org/10.1037/0033-295X.85.3.207>
- Nosofsky, R. M. (2011). The generalized context model: An exemplar model of classification. In E. M. Pothos & A. J. Wills (Eds.), *Formal approaches in categorization* (pp. 18–39). Cambridge: Cambridge University Press. <https://doi.org/10.1017/cbo9780511921322.002>
- Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, 144, 190–206. <https://doi.org/10.1016/j.actpsy.2013.06.003>
- Ossadnik, W., Wilmmsmann, D., & Niemann, B. (2013). Experimental evidence on case-based decision theory. *Theory and Decision*, 75(2), 211–232. <https://doi.org/10.1007/s11238-012-9333-4>

- Pachur, T., & Bröder, A. (2013). Judgment: A cognitive processing perspective. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(6), 665–681. <https://doi.org/10.1002/wcs.1259>
- Pachur, T., & Olsson, H. (2012). Type of learning task impacts performance and strategy selection in decision making. *Cognitive Psychology*, 65(2), 207–240. <https://doi.org/10.1016/j.cogpsych.2012.03.003>
- Pärnamets, P., Johansson, R., Gidlöf, K., & Wallin, A. (2016). How information Availability interacts with visual attention during judgment and decision tasks. *Journal of Behavioral Decision Making*, 29(2–3), 218–231. <https://doi.org/10.1002/bdm.1902>
- Peterson, M. S., & Beck, M. R. (2011). Eye movements and memory. In P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The oxford handbook of eye movements* (pp. 579–592). Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199539789.013.0032>
- Platzer, C., & Bröder, A. (2013). When the rule is ruled out: Exemplars and rules in decisions from memory. *Journal of Behavioral Decision Making*, 26(5), 429–441. <https://doi.org/10.1002/bdm.1776>
- Platzer, C., Bröder, A., & Heck, D. W. (2014). Deciding with the eye: How the visually manipulated accessibility of information in memory influences decision behavior. *Memory and Cognition*, 42(4), 595–608. <https://doi.org/10.3758/s13421-013-0380-z>
- R Core Team . (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. <http://www.R-project.org>
- Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62(8), 1457–1506. <https://doi.org/10.1080/17470210902816461>
- Renkewitz, F., & Jahn, G. (2012). Memory indexing: A novel method for tracing memory processes in complex cognitive tasks. *Journal of Experimental Psychology: Learning Memory and Cognition*, 38(6), 1622–1639. <https://doi.org/10.1037/a0028073>
- Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: Eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *Journal of Experimental Psychology: General*, 133(1), 46–62. <https://doi.org/10.1037/0096-3445.133.1.46>
- Richardson, D. C., & Spivey, M. J. (2000). Representation, space and hollywood squares: Looking at things that aren't there anymore. *Cognition*, 76(3), 269–295. [https://doi.org/10.1016/S0010-0277\(00\)00084-6](https://doi.org/10.1016/S0010-0277(00)00084-6)
- Rosner, A., Schaffner, M., & von Helversen, B. (2022). When the eyes have it and when not: How multiple sources of activation combine to guide eye movements during multiattribute decision making. *Journal of Experimental Psychology: General*, 151(6), 1394–1418. <https://doi.org/10.1037/xge0000833>
- Rosner, A., & von Helversen, B. (2019). Memory shapes judgments: Tracing how memory biases judgments by inducing the retrieval of exemplars. *Cognition*, 190, 165–169. <https://doi.org/10.1016/j.cognition.2019.05.004>
- Rozin, P., & Todd, P. M. (2015). The evolutionary psychology of food intake and choice. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 1–23). Hoboken, NJ: Wiley. <https://doi.org/10.1002/9781119125563.evpsych106>
- Scheibehenne, B., von Helversen, B., & Rieskamp, J. (2015). Different strategies for evaluating consumer products: Attribute- and exemplar-based approaches compared. *Journal of Economic Psychology*, 46, 39–50. <https://doi.org/10.1016/j.joep.2014.11.006>
- Schlegelmilch, R., Wills, A. J., & von Helversen, B. (2022). A cognitive category-learning model of rule abstraction, attention learning, and contextual modulation. *Psychological Review*, 129(6), 1211–1248. <https://doi.org/10.1037/rev0000321>
- Schoemann, M., Schulte-Mecklenbeck, M., Renkewitz, F., & Scherbaum, S. (2019). Forward inference in risky choice: Mapping gaze and decision processes. *Journal of Behavioral Decision Making*, 32(5), 521–535. <https://doi.org/10.1002/bdm.2129>
- Scholz, A., Krems, J. F., & Jahn, G. (2017). Watching diagnoses develop: Eye movements reveal symptom processing during diagnostic reasoning. *Psychonomic Bulletin and Review*, 24(5), 1398–1412. <https://doi.org/10.3758/s13423-017-1294-8>
- Scholz, A., Mehlhorn, K., & Krems, J. F. (2016). Listen up, eye movements play a role in verbal memory retrieval. *Psychological Research*, 80(1), 149–158. <https://doi.org/10.1007/s00426-014-0639-4>
- Scholz, A., von Helversen, B., & Rieskamp, J. (2015). Eye movements reveal memory processes during similarity- and rule-based decision making. *Cognition*, 136, 228–246. <https://doi.org/10.1016/j.cognition.2014.11.019>
- Schulte-Mecklenbeck, M., Johnson, J. G., Böckenholt, U., Goldstein, D. G., Russo, J. E., Sullivan, N. J., & Willemsen, M. C. (2017). Process-tracing methods in decision making: On growing up in the 70s. *Current Directions in Psychological Science*, 26(5), 442–450. <https://doi.org/10.1177/0963721417708229>
- Schulte-Mecklenbeck, M., Kühberger, A., Gagl, B., & Hutzler, F. (2017). Inducing thought processes: Bringing process measures and cognitive processes closer together. *Journal of Behavioral Decision Making*, 30, 1001–1013. <https://doi.org/10.1002/bdm.2007>
- Shadlen, M. N., & Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron*, 90(5), 927–939. <https://doi.org/10.1016/j.neuron.2016.04.036>
- Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2022). *afex: Analysis of factorial experiments* (R package version 1.1-1). <https://CRAN.R-project.org/package=afex>
- Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, 53(1), 1–26. <https://doi.org/10.1016/j.cogpsych.2005.10.003>
- Theeuwes, J. (2018). Visual selection: Usually fast and automatic; seldom slow and volitional. *Journal of Cognition*, 1(1), 29. <https://doi.org/10.5334/joc.13>
- van Horen, F., & Pieters, R. (2012). Consumer evaluation of copycat brands: The effect of imitation type. *International Journal of Research in Marketing*, 29(3), 246–255. <https://doi.org/10.1016/j.ijresmar.2012.04.001>

- Verosky, S. C., & Todorov, A. (2010). Differential neural responses to faces physically similar to the self as a function of their valence. *NeuroImage*, 49(2), 1690–1698. <https://doi.org/10.1016/j.neuroimage.2009.10.017>
- Verplanken, B., & Orbell, S. (2022). Attitudes, habits, and behavior change. *Annual Review of Psychology*, 73, 327. <https://doi.org/10.1146/annurev-psych-020821-011744>
- von Helversen, B., Herzog, S. M., & Rieskamp, J. (2014). Haunted by a Doppelgänger. *Experimental Psychology*, 61(1), 12–22. <https://doi.org/10.1027/1618-3169/a000221>
- von Helversen, B., Karlsson, L., Mata, R., & Wilke, A. (2013). Why does cue polarity information provide benefits in inference problems? The role of strategy selection and knowledge of cue importance. *Acta Psychologica*, 144(1), 73–82. <https://doi.org/10.1016/j.actpsy.2013.05.007>
- von Helversen, B., & Rieskamp, J. (2009). Models of quantitative estimations: Rule-based and exemplar-based processes compared. *Journal of Experimental Psychology: Learning Memory and Cognition*, 35(4), 867–889. <https://doi.org/10.1037/a0015501>
- Wantz, A. L., Martarelli, C. S., & Mast, F. W. (2016). When looking back to nothing goes back to nothing. *Cognitive Processing*, 17(1), 105–114. <https://doi.org/10.1007/s10339-015-0741-6>
- Warlop, L., & Alba, J. W. (2004). Sincere flattery: Trade-dress imitation and consumer choice. *Journal of Consumer Psychology*, 14 (1–2), 21–27. https://doi.org/10.1207/s15327663jcp1401&2_4
- Weber, E. U., & Johnson, E. J. (2006). Constructing preferences from memory. In S. Lichtenstein & P. Slovic (Eds.), *The construction of preference* (pp. 397–410). Cambridge: Cambridge University Press. <https://doi.org/10.1017/cbo9780511618031.022>
- Weilbächer, R. A., Krajbich, I., Rieskamp, J., & Gluth, S. (2021). The influence of visual attention on memory-based preferential choice. *Cognition*, 215, 104804. <https://doi.org/10.1016/j.cognition.2021.104804>
- Wirebring, L. K., Stillesjö, S., Eriksson, J., Juslin, P., & Nyberg, L. (2018). A similarity-based process for human judgment in the parietal cortex. *Frontiers in Human Neuroscience*, 12, 481. <https://doi.org/10.3389/fnhum.2018.00481>
- Wynn, J. S., Shen, K., & Ryan, J. D. (2019). Eye movements actively reinstate spatiotemporal mnemonic content. *Vision*, 3(21), 1–19. <https://doi.org/10.3390/vision3020021>
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35(2), 151–175. <https://doi.org/10.1037/0003-066X.35.2.151>
- Zajonc, R. B. (2000). Feeling and thinking: Closing the debate over the independence of affect. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 31–58). Cambridge: Cambridge University Press.

Appendix A

Table A1. Item structure.

Item no.	Cue 1	Cue 2	Cue 3	Cue 4	Similarity
1	Apple juice	Raspberries	Banana	Beets	4
2	Vanilla soymilk	Strawberries	Blueberries	Oatmeal	4
3	Mixed fruit juice	Pineapple	Mango	Frozen yogurt	4
4	Orange juice	Carrot	Lemon	Ginger	4
5	Apple juice	Raspberries	Oatmeal	Beets	3
6	Apple juice	Pineapple	Banana	Beets	3
7	Orange juice	Raspberries	Banana	Beets	3
8	Apple juice	Strawberries	Blueberries	Oatmeal	3
9	Vanilla soymilk	Mango	Blueberries	Oatmeal	3
10	Vanilla soymilk	Strawberries	Blueberries	Ginger	3
11	Mixed fruit juice	Pineapple	Mango	Raspberries	3
12	Mixed fruit juice	Pineapple	Strawberries	Frozen yogurt	3
13	Orange juice	Pineapple	Mango	Frozen yogurt	3
14	Orange juice	Carrot	Lemon	Banana	3
15	Orange juice	Carrot	Blueberries	Ginger	3
16	Orange juice	Pineapple	Lemon	Ginger	3
17	Apple juice	Raspberries	Strawberries	Frozen yogurt	2
18	Apple juice	Oatmeal	Banana	Carrot	2
19	Orange juice	Raspberries	Banana	Frozen yogurt	2

Table A1. (Continued).

Item no.	Cue 1	Cue 2	Cue 3	Cue 4	Similarity
20	<i>Vanilla soymilk</i>	<i>Strawberries</i>	Raspberries	Mango	2
21	<i>Vanilla soymilk</i>	Beets	Carrot	<i>Oatmeal</i>	2
22	Orange juice	<i>Strawberries</i>	Frozen yogurt	<i>Oatmeal</i>	2
23	<i>Mixed fruit juice</i>	Beets	<i>Mango</i>	Oatmeal	2
24	<i>Mixed fruit juice</i>	Banana	Ginger	<i>Frozen yogurt</i>	2
25	Vanilla soymilk	Lemon	<i>Mango</i>	<i>Frozen yogurt</i>	2
26	Vanilla soymilk	<i>Carrot</i>	<i>Lemon</i>	Banana	2
27	Apple juice	<i>Carrot</i>	Pineapple	<i>Ginger</i>	2
28	Mixed fruit juice	Blueberries	<i>Lemon</i>	<i>Ginger</i>	2
29	<i>Apple juice</i>	Oatmeal	Mango	Carrot	1
30	Vanilla soymilk	Pineapple	Lemon	<i>Beets</i>	1
31	Mixed fruit juice	<i>Raspberries</i>	Strawberries	Carrot	1
32	Mixed fruit juice	Lemon	<i>Blueberries</i>	Raspberries	1
33	Apple juice	<i>Strawberries</i>	Frozen yogurt	Carrot	1
34	Mixed fruit juice	Banana	<i>Blueberries</i>	Ginger	1
35	Apple juice	<i>Pineapple</i>	Strawberries	Lemon	1
36	<i>Mixed fruit juice</i>	Beets	Blueberries	Ginger	1
37	Orange juice	Beets	<i>Mango</i>	Oatmeal	1
38	Vanilla soymilk	Beets	Pineapple	<i>Ginger</i>	1
39	<i>Orange juice</i>	Banana	Blueberries	Frozen yogurt	1
40	Vanilla soymilk	Raspberries	<i>Lemon</i>	Mango	1

Note: Italic indicates exemplar items. Similarity is based on the number of matches between an item and the most similar exemplar. Items with similarity = 1 were randomly assigned to exemplars before the data collection began.

Appendix B

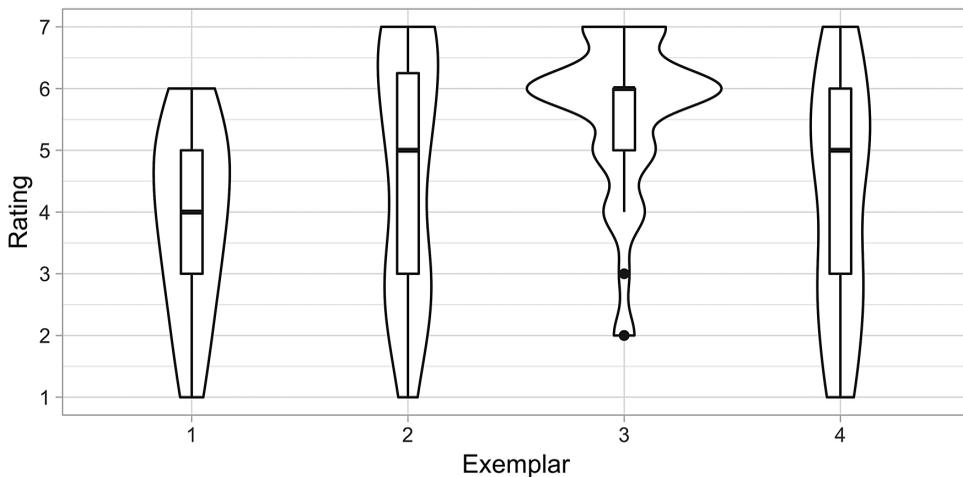


Figure B1. Box plots with density distributions for the fourth exemplar rating during the training phase of participants in the preference condition of Experiment 1. Exemplars consisted of the following ingredients: Exemplar 1: apple juice, raspberries, banana, beets; Exemplar 2: vanilla soymilk, strawberries, blueberries, oatmeal; Exemplar 3: mixed fruit juice, pineapple, mango, frozen yogurt, Exemplar 4: orange juice, carrot, lemon, ginger.

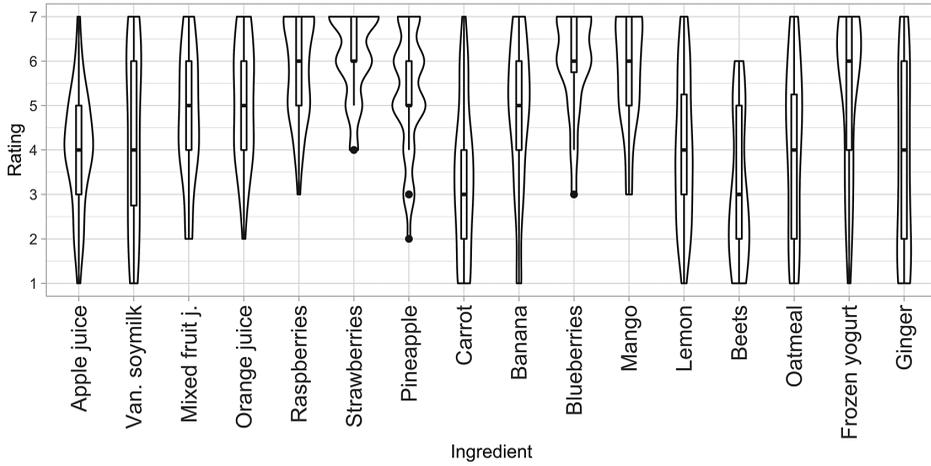


Figure B2. Box plots with density distributions for the ingredient rating of participants in the preference condition of Experiment 1. Van = Vanilla; j = juice.

Appendix C

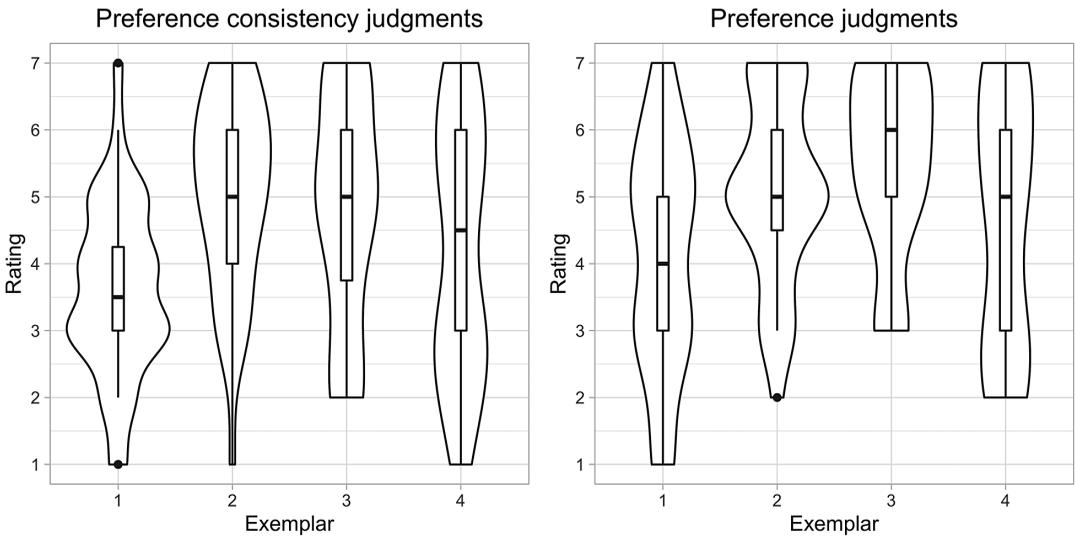


Figure C1. Box plots with density distributions for the fourth exemplar rating during the training phase of Experiment 2. Exemplars consisted of the following ingredients: Exemplar 1: apple juice, raspberries, banana, beets; Exemplar 2: vanilla soymilk, strawberries, blueberries, oatmeal; Exemplar 3: mixed fruit juice, pineapple, mango, frozen yogurt, Exemplar 4: orange juice, carrot, lemon, ginger.

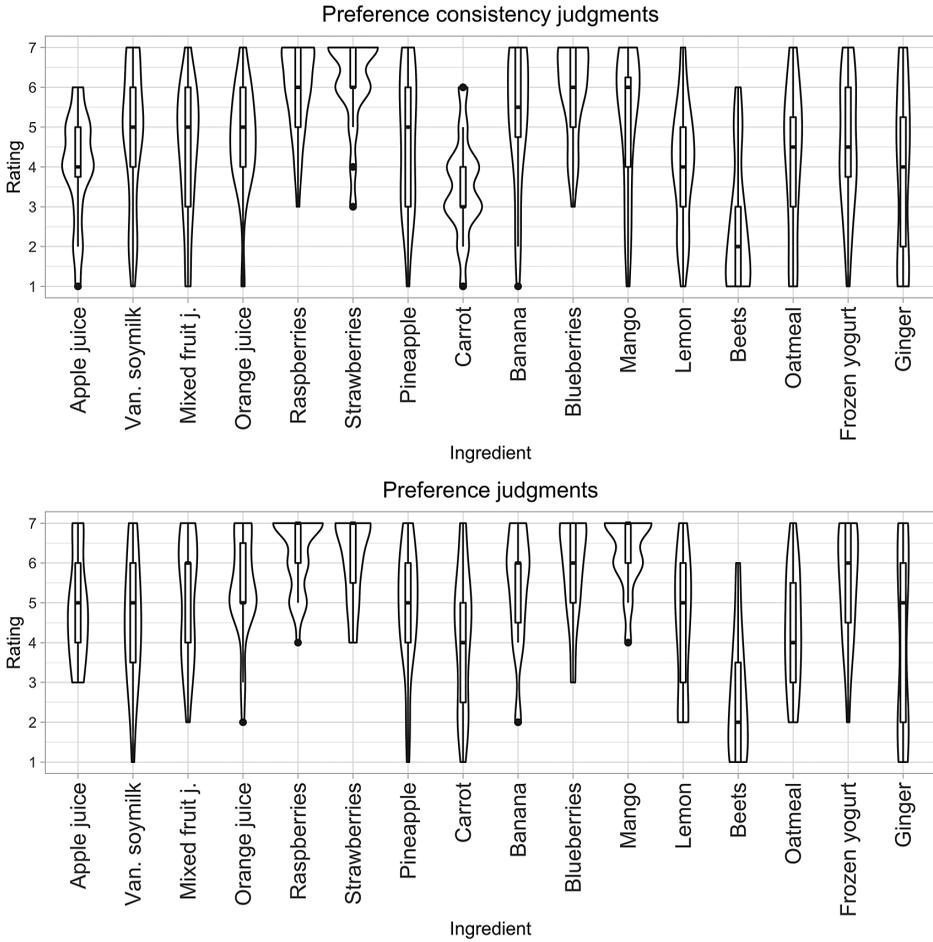


Figure C2. Box plots with density distributions for the ingredient rating of participants in the reference conditions of Experiment 2. Van = Vanilla; j = juice.