Journal: Visual Cognition, Vol.22, No. 1, pp.11-29, 2014.

Title: Gaze bias during visual preference judgments: effects of stimulus category and decision instructions

Authors: Takashi Mitsuda 1) and Mackenzie G. Glaholt 2)

Author 1) Affiliation: Department of Human and Computer Intelligence, Ritsumeikan University, Shiga 525-8577, Japan

E-mail: mitsuda@ci.ritsumei.ac.jp **Telephone/Fax:** +81-77-561-5068

Author 2) Affiliation: Defence Research and Development Canada, Toronto, Canada

E-mail: mackenzie.glaholt@gmail.com

Corresponding author: Takashi Mitsuda

Keywords: Eye movements; Preference; Gaze bias; Decision making

To cite this article: Takashi Mitsuda & Mackenzie G. Glaholt (2014) Gaze bias during visual preference judgements: Effects of stimulus category and decision instructions, Visual Cognition, 22:1, 11-29, DOI: 10.1080/13506285.2014.881447

To link to this article: http://dx.doi.org/10.1080/13506285.2014.881447

Abstract

Prior research has demonstrated that during two-alternative decision making, gaze is biased towards the alternative that is eventually chosen. The Gaze Cascade model proposed by Shimojo, Simion, Shimojo, & Scheier (2003) predicts a larger bias for decisions requiring one to choose the item that is liked the most versus decisions that require one to choose the item that is disliked most. More recently, Park, Shimojo and Shimojo (2010) showed that preference formation operates differently during decisions among faces and scenes, which suggests that gaze bias might differ depending on whether the decision stimuli are faces or scenes. In the present study we tested these two hypotheses in a within-subject design. Eye movements were monitored while participants (n=48) made two-alternative Like or Dislike decisions among pairs of faces or scenes. We found remarkably little influence of stimulus type on gaze bias for either decision task, which disconfirms the hypothesis that gaze bias operate differently for faces than scenes. In contrast, we found that gaze bias was stronger for Like decisions than Dislike decisions. To further account for this effect we examined the decision time course, which revealed that this task effect is primarily related to biases in the placement, and duration, of the final dwell prior to response, though there was evidence that the bias began early for Like decisions. Implications for mechanisms of gaze allocation during multi-alternative decision making are discussed.

Consumers often make decisions that require them to consider a set of possible alternatives and choose one or more for purchase. This might occur in a retail store setting where the consumer browses products on store shelves, though it is increasingly possible to make purchasing decisions online and survey the decision alternatives with a web browser. Consumer preference decisions can be thought of as instances of a general class of multi-alternative decisions that has received decades of interest in psychology, cognitive science, and economics, resulting in a rich domain of theory (for a review see Gigerenzer and Gaissmaier, 2011; see also Reidl, Brandstätter & Roithmayr, 2008) and mathematical modeling (for a review, see Tsetsos, Usher, and Chater, 2010). In addition to the modeling of decision outcomes (e.g. choices), prior research has considered the time course of the decision process (e.g. Roe, Busemeyer, & Townsend, 2001; Usher & McClelland, 2001), and process-tracing methodologies have been developed in order to provide empirical data about the time course of decision making (for recent reviews see Glaholt & Reingold, 2011; Russo, 2011). In particular, these methods seek to measure the overt sampling of information about decision alternatives by the decision maker. Visual sampling is a primary means by which information about decision alternatives is gathered, and hence eye movement monitoring (e.g. eye-tracking) has the potential to be an important empirical tool for the study of decision making.

A rapidly growing body of recent research has used eye-tracking to monitor visual sampling behavior during multi-alternative visual preference decisions (e.g. Bird, Lauwereyns, & Crawford, 2012; Glaholt & Reingold, 2009a, 2009b, 2011; Glaholt, Wu & Reingold, 2010; Krajbich, Armel, & Rangel, 2010; Krajbich & Rangel, 2011; Liao & Shimojo, 2011; Nittono & Wada, 2009; Schotter, Berry, McKenzie, & Rayner, 2010; Schotter, Gerety, & Rayner, 2012; Shimojo, Simion, Shimojo & Scheier, 2003; Simion & Shimojo, 2006, 2007; for a review see Orquin and Loose, 2013). Interest in this domain was spurred by Shimojo, et al. (2003), who monitored eye movements while decision makers viewed pairs of human faces and decided which one was more attractive. The authors introduced an analysis of the likelihood of participants' gaze being directed to the chosen alternative over the interval prior to the response (gaze likelihood analysis). This revealed that gaze was initially distributed evenly between the two stimuli, but prior to the response the gaze likelihood became biased towards the face that was chosen. The discovery of this 'gaze bias' prompted Shimojo and colleagues to put forward a model of preference formation (the "Gaze Cascade" model) in which preferential looking and the mere exposure effect form a positive feedback loop that leads to a decision outcome. Under this proposed mechanism, exposure to an item increases preference, and preference increases the likelihood of looking at that item, which in turn increases exposure.

There are two related predictions that can be derived from this model and they are the focus of the present paper. First, because the Gaze Cascade model depends on a positive effect of exposure on preference formation, it would predict a reduction in the gaze bias under conditions where exposure decreases preference, and recent studies have been suggested that when participants made a preference decision for natural scenes exposure can decrease rather than increase preference (Liao, Yeh, & Shimojo, 2011; Park, Shimojo, & Shimojo, 2010). Second, the model hypothesizes that the gaze bias is related to preference-specific mechanisms (preferential looking and the mere exposure effect), and hence the gaze bias effect should be unique to, or uniquely strong for, preference decisions compared to other decisions. In the following section we expand upon these predictions in the context of prior findings, and subsequently we introduce the design of the present study.

The prediction regarding the preference-specificity of the gaze bias has been tested in several studies, resulting in a mixed pattern of findings (Glaholt & Reingold, 2009a, 2009b, 2011; Nittono & Wada, 2009; Schotter, et al., 2010; Shimojo et al., 2003; Simion & Shimojo, 2006). In support of the model's prediction, Shimojo and colleagues

(Shimojo et al., 2003, Simion & Shimojo, 2006) demonstrated a larger gaze likelihood effect for preference decisions than for other decisions (face roundness, dislike). However, Glaholt & Reingold (2009a, 2009b, 2011) compared preference decisions with two non-preference decisions (recency and typicality decisions) and found that the gaze bias effect was nearly identical for all decision instructions tested. Similarly, Nittono and Wada (2009) found no differences in gaze bias as a function of decision instructions (dislike, brightness decisions). On the other hand, Schotter et al. (2010) did observe differences in the magnitude of the gaze bias as a function of decision instructions (like decision instructions showed a greater bias than dislike and recency instructions). These various findings point to a potentially complex relationship between gaze bias and decision task instructions. One factor that might have contributed to these differences across studies relates to the stimuli used during the decision tasks. Shimojo et al. (2003) originally used faces and randomly generated geometrical patterns, while Nittono and Wada (2009) used graphic patterns, Glaholt and Reingold (2009a, 2009b, 2011) used stimuli that fall under the category of scenes, and Schotter et al. (2010) used stimuli composed of scenes and portraits.

Recent work by Shimojo and colleagues (Liao & Shimojo, 2011; Liao, et al., 2011; Park, et al., 2010) has suggested that stimulus category might be particularly important factor in the context of the gaze bias. In particular, Park et al. (2010) and Liao et al. (2011) developed a two-alternative forced choice paradigm where participants made a preference decision between a novel stimulus and stimulus they had seen before. Stimulus category was manipulated between trials. For decisions among faces, preference tended to be biased towards the repeated face (familiarity preference) while for scenes, preference tended to be biased towards the novel scene (novelty preference). This dissociation in the effect of exposure on preference as a function of stimulus type is particularly important in the context of the present study because it provides an opportunity to test our prediction derived from the Gaze Cascade model. To reiterate, the Gaze Cascade mechanism describes a positive feedback loop where gaze produces exposure to a stimulus which increases the preference for that stimulus, and preference increases the likelihood of looking at that item, which in turn increases exposure. Because a positive effect of exposure on preference is integral to this mechanism, it predicts a reduced gaze bias effect for conditions for which exposure decreases preference. Based on the work of Park et al. (2010) and Liao et al. (2011), a manipulation of stimulus category between faces and scenes should provide this contrast. To our knowledge the only study to date that measured gaze bias and that featured a direct comparison between faces and scenes is a study by Liao & Shimojo

(2011). In this study, the authors applied a gaze likelihood analysis (e.g. Shimojo et al., 2003) and reported that the gaze bias prior to the response did not differ as a function of stimulus category, while there was some evidence of stimulus-specificity in looking behavior early along the decision time course.

Indeed, the manifestation of the gaze bias over the decision time course has been a major interest in prior research. The gaze bias phenomenon was originally expressed as an increase in the likelihood of looking at the chosen item in the interval just prior to the response. Concerns have been raised (Glaholt & Reingold, 2009a, 2011; Nittono and Wada, 2009), however, that the gaze likelihood analysis that was used by Shimojo et al. (2003) failed to capture the discrete nature of gaze behavior over the decision time course. More specifically it is unclear whether the bias in gaze likelihood is driven by a tendency to look at the chosen item last (i.e. just prior to the response) or if gaze behavior is also biased earlier in the decision time course. To address this, Glaholt and Reingold (2009a) introduced an analysis that divided the time course into a sequence of dwells, where each dwell was defined as a run of one or more consecutive fixations on a decision alternative. This analysis revealed that for eight-alternative decisions, the bias toward the chosen item was present from the very first dwell in the trial. Similarly, Schotter et al. (2010) and Schotter et al. (2012) reported a bias in the dwell duration of the first encounters with the decision alternatives during two-alternative decisions. These findings tended to rule-out purely response-related explanations for the gaze bias phenomenon, and also underlined the importance of considering the detailed time course of eye movements during multi-alternative decisions.

In summary, the present study sought to test two hypotheses derived from the Gaze Cascade model concerning the stimulus- and task-specificity of the gaze bias phenomenon. The first hypothesis is that the gaze bias would be reduced under conditions where exposure decreases preference (scene stimuli) compared to conditions where exposure increases preference (faces). The second hypothesis is that the gaze bias effect should be stronger for preference decisions than other decisions. Accordingly, we conducted a within-subject design crossing decision task (like vs. dislike instructions) and stimulus type (faces vs. scenes). Based on the predictions derived from the Gaze Cascade model we expected an interaction where the gaze bias would be most pronounced for face stimuli under the like instructions, and smallest for scene stimuli under dislike instructions.

Beyond the specific predictions of the Gaze Cascade model, any effects of decision task and stimulus type that we observe will provide strong constraints for theoretical modeling efforts in this domain. For example, the preference-specific account by

Schotter et al., (2010) and Schotter et al. (2012) was motivated by the finding of an early difference in the gaze bias time course between like and dislike task instructions. Other work has portrayed the gaze bias as a general characteristic of visual decision making tasks. For example, Glaholt and colleagues (Glaholt & Reingold, 2009a, 2009b, 2011) suggested that the gaze bias reflects the selective encoding of task relevant information, regardless of the specific decision task instructions. Schotter et al. (2010, 2012) also explained their data in terms of the additive effect of the selective encoding and the preferential looking. Similarly, work by Krajbich and colleagues (Krajbich, et al., 2010; Krajbich & Rangel, 2011) provided a general computational account of the gaze bias phenomenon as component of the accumulation of decision evidence towards a threshold for response. Therefore, while the central hypotheses of the present study were derived from the Gaze Cascade model by Shimojo et al. (2003), we expect the present study to provide important empirical evidence that will guide and differentiate modeling efforts in this domain.

Method

Participants

Forty eight male students at Ritsumeikan University (Mean age: 22.0 years, age range: 18–24 years) participated in the experiment, and received 800JPY for their participation. All participants signed informed consent forms and had normal or corrected-to-normal vision.

Apparatus

The participants were seated 57 cm away from a 19-in monitor. The monitor was set to a resolution of 1024 × 768 and a refresh rate of 60 Hz. Eye movements were recorded using an Eyelink II head-mounted video-based eye tracker (SR Research Ltd.) at 250 Hz. Each stimulus (two images side by side) subtended a visual angle of 30° (H) × 15° (V).

Materials and Procedure

The stimuli consisted of 60 pairs of female faces and 60 pairs of scenes. For the present experiment, our definition of the 'scene' category included landscapes, animals, flowers, and architecture (see Figure 1, panel a). The pairs of female faces consist of 120 faces of female models for hair styles collected from the Internet (see Figure 1, panel b). Similar attributes of the faces (e.g., hairstyle, makeup, age) were taken into consideration when assembling pairs of images. The background for all of the face images was made white. The pairs of scenes consist of 120 photographs collected from the Internet. Scenes within each pair included the same content (e.g., both sunsets, both rabbits, both Cherry blossoms, both churches). All of the images were in color.

In the experiment, participants were asked to select the image they liked more (Like task) or the image they disliked more (Dislike task) for each pair of images. The instruction was made orally in Japanese ("Sukina hou wo erande kudasai" for the Like task and "Kiraina hou wo erande kudasai" for the Dislike task). Each subject performed four tasks (like-face task, dislike-face task, like-scene task, dislike-scene task). The order of tasks was counterbalanced across participants. Each participant saw 30 pairs of images in a task. The pairs of images used for each task and the order of presenting pairs of images were randomized across participants. No participant saw any image more than once.

The eye tracker was calibrated and validated before each block of recording. The trial began with a fixation on a gray dot in the center of the screen. After fixating on the dot, a pair of images appeared on the screen. The participants were instructed to view the images freely and press the right arrow key when they selected the right image or the left arrow key when they selected the left image. The images disappeared once they pressed an arrow key, and the next trial began with a central fixation. The four blocks were recorded on the same day.

Results

Fixation events were identified from the raw eye movement record using SR Research Eyelink Dataviewer. Throughout the analyses we consider only eye movement data (gaze samples or fixation events) that occurred following the presentation of the two-alternative decision display (see Figure 1) and prior to the button response that indicated the two-alternative forced-choice response. Fixations that spanned these boundaries were truncated. We defined a dwell (as in Glaholt & Reingold, 2009a; Schotter et al., 2010) as one or more consecutive fixations on a single image, and a dwell terminated when the participant fixated the other image. When consecutive dwells interleaving fixations out of images are on a single image, these dwells were treated as a single dwell. Dwell duration was defined as the cumulative duration of the fixations that composed the dwell, and the duration of fixations that fell outside either image was excluded. The analyses are presented in two sections: first we considered global measures of gaze bias; second, we examined the time course of gaze bias over the interval prior to the response.

Global measures of gaze bias

In order to measure the overall pattern of gaze bias, we computed total dwell duration, number of dwells, and mean dwell duration and analyzed them in a 2 x 2 x 2 repeated measures ANOVA that crossed Stimulus (face vs. scene), Decision (like vs. dislike) and Item (chosen item vs. other). As expected, we observed a gaze bias in total dwell duration (see Figure 2, panel a), with longer duration on the chosen item compared to the not chosen item (F(1.47) =103.8, $p < 0.001$). This overall bias was associated with an increased number of dwells, (F(1,47)=128.9, $p < 0.001$) and an increase in mean dwell duration (F(1,47) =19.2, $p < 0.001$) for the chosen item compared to the other item (see Figure 2, panels b and c). There were main effects of Stimulus type where, compared to faces, decisions among scenes received more total dwell time (F(1,47) =19.0, $p < 0.001$), a greater number of dwells (F(1,47) =33.9, $p < 0.001$), and longer mean dwell duration (F(1,47) =4.6, $p < 0.05$). Importantly, however, for each of these measures the Stimulus factor did not interact significantly with Item (all Fs <2.21, ps > 0.14), or with Decision and Item (all Fs < 1), indicating that the gaze bias captured by these global measures was not sensitive to the stimulus category of the decision alternatives. There were also strong effects of the Decision factor, where Dislike decisions has longer total dwell duration (F(1,47) =8.9, $p < 0.01$) and more dwells (F(1,47) =11.9, $p < 0.01$) than Like decisions. More importantly, there was a significant interaction between Decision and Item where the gaze bias was larger for the Like decision than the Dislike decision in total dwell duration (F(1,47) =13.4, $p < 0.001$), number of dwells ($F(1,47) = 8.1$, $p < 0.01$) and in mean dwell duration ($F(1,47) = 6.6$, $p < 0.05$). In order to highlight this pattern of effects, we computed a proportional gaze bias measure which is the total dwell duration on the chosen item was expressed as a proportion of the total dwell duration. This measure (shown in Figure 2, panel d) highlights the effect of Decision on gaze bias as well as its insensitivity to the Stimulus factor.

Time course of gaze bias

The Gaze Cascade model was originally proposed to account for an increase in the likelihood of looking at the chosen item over the interval just prior to the response. We applied the gaze likelihood analysis of Shimojo et al. (2003), and subsequently we conducted a dwell sequence analysis (Glaholt & Reingold, 2009a) which examined the entire time course, including the early part of the decision interval.

For the gaze likelihood analysis we plotted, for each eye movement sample prior to the response, the probability of gaze being directed to the chosen item. This measure is computed across trials for a given task (i.e. Stimulus and Decision type), separately for each subject. For each time point prior to the response, the gaze likelihood value corresponds to the number of trials in which gaze was directed to the chosen item at that time point divided by the number of trials for which gaze was directed to either item at that time point (therefore chance = 0.5). The gaze likelihood curves for each Stimulus and Decision type are displayed in Figure 3. Dotted lines depict 95% confidence intervals about the mean. As can be seen in the Figure, a bias in gaze likelihood (described as a 'gaze cascade' effect by Shimojo et al., 2003) was evident in all conditions though there were some apparent differences. Figure 4 shows the final gaze likelihood (i.e. probability of looking at the chosen item last) for each Stimulus and Decision type. There was a strong bias in the placement of the last dwell to the chosen item (for comparison against chance, all *t*s >7.0, all *p*s < 0.001). This variable showed similar pattern to that which was observed in the overall gaze bias. More specifically, there was a strong effect of Decision ($F(1,47) = 25.6$, $p < 0.001$) with the Like decision showing a greater probability of looking at the chosen item last than the Dislike decision. However there was also a significant effect of Stimulus type ($F(1,47) = 6.9$, $p < 0.05$), where scenes had a higher probability of having the last dwell on the chosen item. The interaction between Stimulus and Decision did not approach significance ($F < 1$).

To provide a more detailed examination of the time course of gaze bias, we conducted a dwell sequence analysis. This analysis was introduced by Glaholt & Reingold (2009a) in order to provide evidence of a gaze bias that is prior to the final dwell in the trial. The reason for this is that the final dwell might be lengthened for reasons other than decision-related processing, such as response latency and other post-decision artifacts (for discussion of this, see Glaholt & Reingold, 2011 and Nittono & Wada, 2009). The gaze likelihood analysis does not partition gaze duration based on dwells, and hence it is difficult to gauge the extent to which the shape of the gaze likelihood curve is influenced by the placement and duration of the last dwell in the trial. This analysis also allowed us to test for an early gaze bias in the very first dwell in the trial. Accordingly we computed the mean duration of dwells on the chosen and not chosen item for dwells occurring at each of three dwell sequence positions going back from the response (Last, Last-1, Last-2) and also in the first dwell position in the trial (First). Last dwell duration and Last-1 dwell duration were computed with the trials that have three dwells at least. Last-2 dwell duration was computed with the trials that have at least four dwells. The dwell sequence plots for each condition are presented in Figure 5. We analyzed each dwell sequence position separately in a 2x2x2 repeated measures ANOVA crossing Stimulus (Face vs. Scene), Decision (Like vs. Dislike), and Item (Chosen vs. Other).

There was a strong gaze bias in the last dwell duration ($F(1,39) = 68.6$, $p < 0.001$). The bias showed significant interactions with Decision and Stimulus, being larger Like decisions than Dislike decisions ($F(1,39) = 6.0$, $p = 0.019$), and for scenes compared to faces ($F(1,39) = 12.4$, $p < 0.001$). There was a gaze bias also in the dwell duration for the penultimate dwell (Last-1) ($F(1,31) = 13.0$, $p = 0.001$). The bias interacted with decision type ($F(1,31) = 9.8$, $p < 0.01$), but not with stimulus category ($F < 1$). One-way repeated measures analysis of variance on the penultimate dwell duration showed a gaze bias for Face Like task ($F(1,39) = 16.0$, $p < 0.001$), and for Scene Like task ($F(1,39) = 10.9$, $p = 0.002$), but not for Face Dislike task ($F < 1$), or for Scene Dislike task ($F < 1$). There was no significant effect of Item in the dwell duration for dwells two prior to the last dwell (Last-2) ($F < 1$) nor were there significant interactions between Item and Decision Type ($F < 1$), or with Stimulus Category ($F < 1$). Hence the increase in gaze bias for Like decisions compared to Dislike decisions in the global measures is not exclusively driven by a bias in the placement and duration of the final dwell, but is also manifested on the penultimate dwell.

In the analysis of first dwell durations (First), there was a significant three-way interaction between Stimulus, Decision, and Item ($F(1,47) = 7.4$, $p < 0.01$) but no other significant main effects or interactions. In order to interpret the three-way interaction, we conducted follow-up one-way ANOVAs testing the effect of Item for each of the 2x2 Stimulus and Decision conditions. None of these analyses yielded significant effects of

Item (all *p*s > 0.05), and as can be seen in Figure 5, the differences between Chosen and Not Chosen items in the First dwell position were very small and variable across conditions, and hence we conclude that despite the significant overall 3-way interaction, the gaze bias in the first dwell position was negligible.

This appears to be inconsistent with the finding by Schotter et al. (2010, 2012) of an early bias that was present for Like decisions but absent for Dislike decisions. However, to be clear, Schotter et al. (2010) looked for a bias in the *first encounter* with the chosen and not chosen item, which should be distinguished from a bias in the first dwell position in the trial. More specifically, first encounter would consider the first and second dwell positions in the trial, which will constitute the first encounters with the chosen and not chosen item. The first dwell position only considers the first dwell in a trial, which will be directed either to the chosen or the not chosen item (hence data for the first dwell position on chosen and not chosen items must be drawn from different trials). We examined the first encounter duration (see Figure 6, panel a) and observed a gaze bias for the first encounter for Face Like Decisions but not for the other conditions. This was reflected in a significant three-way interaction between Stimulus, Decision, and Item ($F(1,47) = 4.5$, $p < 0.05$). One-way repeated measures ANOVA showed that the first encounter duration had a bias for Face-Like task ($F(1,47) = 13.9$, p

< 0.001), but not for other tasks (all *F*s <1.3, *p*s >0.28). This partly replicates prior findings of Schotter et al. (2010). However, in order to better understand the cause of this effect we divided trials according to whether there were two, three, or four, or five or more dwells in the trial. The relative proportions of trials with each number of dwells can be seen in Figure 6, panel b, and first encounter duration, broken down by each number of dwell setting are presented in Figure 6, panels c, d, e, and f. Examination of the Figure reveals that the difference in first encounter dwell duration for the chosen and other items depended strongly on the number of dwells in the trial. One-way repeated measures ANOVA showed that the first encounter dwell duration had a significant bias for each number of dwells (all *p*s<0.001), though the direction of the effect differed depending on the number of dwells. In particular, trials with either two dwells or four dwells showed a large bias in the first encounter duration, while trials with 3, 5 or more showed a reversed bias.

This strong variation in the magnitude and direction of first encounter gaze bias as a function of trial length is especially important to consider given the finding from global analyses that both Stimulus and Decision had a significant effect on the number of dwells in a trial. Indeed, the proportion of trials with each number of dwells for each condition (see Figure 6, panel b) differed across conditions. For example, the Face Like task had a larger proportion of trials with 2 and 3 dwells than the Scene Like task. Given that the first encounter bias is especially pronounced in trials with only 2 dwells, the three-way interaction between Stimulus, Decision, and Item seen in the overall first encounter measure might be related to the uneven proportions of short duration trials across conditions.

Another important factor to consider with regards to the first encounter bias is that, as can be seen in the average dwell sequence curves in Figure 5, regardless of condition, the average dwell duration changes greatly over the decision time course. In particular, the first dwell and last dwells are relatively short, while the middle dwells (e.g. Last-2, Last-1 from Figure 5) are relatively long. First encounter dwells in a two-alternative forced choice trial are necessarily the first two dwells, and hence the magnitude of the gaze bias in first encounter might be expected to depend strongly on whether first dwell is directed to the chosen or the not-chosen item. For example, in the trials with an odd number of dwells, the first encounter dwell on the chosen image tended to be at the first dwell position because the last dwell tended to be on the chosen item. In the same way, the first encounter dwell on the not-chosen item in the trials tended to be at the second dwell position. Therefore, the gaze bias for first encounter dwells might reflect the difference in dwell duration between the first dwell and the second dwell.

In order to address this concern, we separated trial sequences based on the identity of the first dwell. First we computed the likelihood of subjects choosing the first item that was viewed. A 2x2 ANOVA crossing Stimulus and Decision revealed no significant effects or interactions (all *F*s < 1), indicating that there was no bias in the placement of the first dwell. We then computed dwell sequence time courses for trials of 2, 3, 4, 5, 6, or 7 dwells in length as a function of whether the first dwell was directed to the chosen or other item. Figure 7 depicts the time courses for each of these sequences relative to the beginning of the trial, collapsing across both Stimulus and Decision factors. The overall shape of these time courses is roughly consistent with dwell sequence time course presented in Figure 5. However, what is clear in these individual time courses is that the final dwell in the trial has a particularly short duration for cases where it was directed to the not chosen item, which would contribute strongly to the overall gaze bias at the final dwell.

Discussion

The present study examined the effect of stimulus category and decision instructions on the gaze bias phenomenon in two-alternative visual decision tasks. Consistent with prior research on the gaze bias phenomenon (e.g. Glaholt & Reingold, 2009a, 2009b, 2011; Schotter, Berry, McKenzie, & Rayner, 2010; Schotter, Gerety, & Rayner, 2012; Shimojo, Simion, Shimojo & Scheier, 2003; Simion & Shimojo, 2006, 2007; for a review see Orquin and Loose, 2013), our data demonstrated a robust bias in overall gaze duration toward the chosen item. Importantly, we found very little evidence for an influence of stimulus category on this overall gaze bias. While decisions among scenes tended to be slightly longer than among faces, the proportional gaze bias was quite similar across both stimulus categories, as was the manifestation of the bias over the decision time course. There were small but significant effects of stimulus type on the probability of looking at the chosen item last, and the magnitude of the bias in the last dwell duration, but in both cases the effect was greater for scenes than for faces. Thus our data failed to support the hypothesis that the gaze bias for decisions among faces would be larger than for decisions among scenes. To reiterate the logic of this hypothesis, the Gaze Cascade mechanism describes a positive feedback loop where gaze produces exposure to a stimulus which increases the preference for that stimulus, and preference increases the likelihood of looking at that item, which in turn increases exposure. This mechanism requires a positive effect of exposure on preference, and hence it would predict a reduced gaze bias effect for conditions in which exposure decreases preference. Park et al. (2010) and Liao et al. (2011), argued that exposure decreases preference for scenes (but increases preference for faces) and hence according to the Gaze Cascade model, the gaze bias should be larger for faces than scenes.

Given that we observed a gaze bias that was highly similar for decisions among faces and scenes, we suggest that any mechanistic account of the gaze bias phenomenon should accommodate different stimulus categories. For example, it has been suggested previously (e.g. Glaholt & Reingold, 2009a, 2009b, 2011; Schotter, et al., 2010, 2012) that the gaze bias might involve the selective encoding of stimulus information where relevant information is receives longer gaze duration and irrelevant information receives less. This account is agnostic with respect to the content of the stimuli. In another recent account of the gaze bias, Krajbich and colleagues (Krajbich, Armel, & Rangel, 2010; Krajbich & Rangel, 2011) have linked eye fixations to the accumulation of evidence in a diffusion process. Under this model, the gaze bias toward the chosen item emerges because the threshold in accumulated evidence for the chosen alternative is most likely to be reached while gaze is directed to the chosen item rather than the other item. This model is also consistent with a gaze bias across face and scene stimuli.

The second goal of the present study concerned the effect of decision instructions on the pattern of gaze bias. We found that consistent with prior reports (Shimojo et al., 2003; Schotter et al., 2010, 2012), Like decisions produced a larger gaze bias than Dislike decisions. This was evident in the gaze likelihood curves and in the proportional overall gaze bias. In addition, this study revealed that while both Like and Dislike decisions exhibit a gaze bias in the final dwell in the trial, only Like decisions exhibited a significant gaze bias prior to the final dwell. This difference in the onset of the gaze bias as a function of decision instructions was not reflected in the gaze likelihood curves, where onset of the gaze bias is roughly the same between Like and Dislike decisions; the gaze likelihood analysis primarily reflected the bias in the placement of the last dwell. As has been discussed in prior work (Glaholt & Reingold, 2009a, Nittono & Wada, 2009), the observation of gaze biases just prior to the response (e.g. in the last dwell position) can be difficult to interpret because of the impending response. In contrast, biases shown earlier in the dwell sequence are not likely to be influenced by the response, and hence can be attributed to decision processes.

We also sought evidence of a gaze bias early in the decision period. Interestingly, the first dwell position did not exhibit a significant bias, either in dwell duration or in the likelihood of viewing the chosen item, under any of the conditions tested. However, we observed a small but significant bias when considering the first encounter with the chosen and not chosen items, but this first encounter bias was only present for the Face Like task and not for any of the other tasks. This finding is partially consistent with the report of Schotter et al. (2010) though in their study the first encounter bias was documented for Like decisions among scenes. Schotter et al. (2010, 2012) suggested that the effect of decision instructions on first encounter gaze bias reflects an early 'liking' effect that increases the gaze bias for Like decisions and decreases it for Dislike decisions, and that may be distinguished from a selective encoding component of the gaze bias that common to Like and Dislike decisions and is manifest later in decision period. However, our results indicated that the first encounter bias depends strongly on the number of dwells in the trial. The higher proportion of trials ending with dwells on the chosen item together with the difference in dwell duration between the first and second dwell enlarges the first encounter bias, especially for the two-dwells trials. Like decisions tended to have a greater proportion of two-dwell trials than Dislike decisions, which is likely to explain the different pattern of gaze biases between decision tasks in average first encounter duration.

Not only the first encounter bias, but the overall gaze bias can be affected by the differences in dwell duration over dwell sequence positions and also by the imbalance of number of trials ending in the chosen item. The time-course analysis separating trials by the number of dwells and by the identity of the first dwell revealed that the last dwell duration ending in the not-chosen item is much shorter than other dwell durations. In contrast, the last dwell duration ending in the chosen item was almost equivalent to the preceding dwell durations. This means that the overall gaze bias in dwell duration is mainly affected by the proportion of trials ending in the not-chosen item and the short dwell duration at the end of trials. The different proportion of trials with a small number of dwells also affected the gaze bias. These findings underline the importance of considering trial composition in terms of dwells and the time course of the gaze bias over those dwells in addition to global measures.

The difference, taken together with the finding of a higher proportional gaze bias for Like decisions than Dislike decisions, and an elevated probability of directing the last dwell to the chosen item, constitute systematic and robust differences between gaze bias under Like and Dislike decision instructions.

These effects of decision instructions are amenable to different processing explanations. The more pronounced bias in gaze likelihood for Like decisions is consistent with the Gaze Cascade mechanism proposed by Shimojo et al. (2003). However, we contend that the model would also have predicted an interaction between decision instructions and stimulus category according to the different preference formation effects found by Park, Shimojo and Shimojo (2010), with the Face Like decisions showing a stronger bias than Scene Like decisions, and this interaction was not a feature of our results. Schotter et al. (2010, 2012) also posited a mechanism that should produce a relatively strong bias for Like decisions. More specifically, they argued for a liking component that increased the first encounter bias for Like decisions but not Dislike decisions, and that may be contrasted with a separate selective encoding effect that emerged later and was common to both Like and Dislike decisions. As was discussed, our data show that the dwell duration bias in first encounters might be an artifact of differing trial lengths under the two task instructions. Nevertheless we found that when considering the time course of the dwell sequence for longer trials (3 or 4 dwells), the bias appeared earlier for Like decisions than Dislike decisions, which is at least broadly consistent with the Schotter et al. (2010, 2012) account.

However, we suggest that the present pattern of findings might also be explained more simply. We note that Like decisions were shorter than Dislike decisions, and hence it is possible that there is a difference in difficulty between Like and Dislike decisions in this task. Importantly, in the present design the same stimuli were considered under Like and Dislike instructions, and hence any difference in difficulty would not be tied to the differences at the level of the stimuli, but instead the difficulty in extracting decision-relevant information under the two task instructions. Such a processing-related difference might exist if participants have more experience or practice at making one decision instruction than another. For example, Like decisions might be made more often under natural choice situations than Dislike decisions, resulting in a more efficient extraction of decision-relevant information a more rapid differentiation between stimuli according to Like decision criteria than Dislike decision criteria. This interpretation predicts an earlier bias for Like than Dislike decisions for trials of a given length (i.e. number of dwells), as well as shorter overall decisions for Like than Dislike decisions, though does not require a fundamental difference in the nature of processing under these decision instructions. Future work might seek to evaluate this hypothesis. In particular, an application of the recent accumulator model by Krajbich et al. (2010, 2012) might provide insight into ways in which these task differences can be accounted for under a single processing architecture.

Another possible explanation of the difference in pattern of gaze bias across Like and Dislike task instructions is that the two decisions involve different strategies. For example, it might be the case that for the Dislike task participants tend to identify the preferred item first and then select the other item. Under this strategy, the Dislike decisions might resemble Like decisions except for, perhaps, a final dwell confirming the item to be selected. This might explain the finding of a significant gaze bias only in the last dwell position for the Dislike task, but such an account might also have predicted a reverse gaze bias prior to the final dwell, which was not a feature of our data. It has also been demonstrated that the adoption of a selection strategy versus a rejection strategy can affect the decision process during two-alternative decisions (Meloy & Russo, 2004). More specifically, selection strategy (e.g. choose which item to keep) produces fluent decisions among appealing alternatives, while a rejection strategy (e.g. choose which item to leave) produces more fluent decisions among unappealing alternatives. While the underlying valence of the stimuli was not manipulated in the present study, it is possible that the observed difference in Like and Dislike tasks might map onto an underlying difference in the fluency with which participants could apply a selection or rejection strategy.

In summary, the present study provided important empirical data on the gaze bias phenomenon in two-alternative decision tasks. We found a highly similar pattern of gaze biases for decisions among faces and scenes and we replicated and extended the finding of robust differences in gaze bias between Like and Dislike task instructions. These findings argue against stimulus-category specific models of the gaze bias while providing support for models that can capture processing differences resulting from decision task instructions. Future work might seek to account for this pattern of results through the quantitative modeling of gaze behavior.

References

Bird, G. D., Lauwereyns, J., & Crawford, M. T. (2012). The role of eye movements in decision making and the prospect of exposure effects. *Vision Research*, 60, 16-21.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62 (1), 451-482.

Glaholt, M. G., & Reingold, E. M. (2009a). The time course of gaze bias in visual decision tasks. *Visual Cognition*, 17 (8), 1228-1243.

Glaholt, M. G., & Reingold, E. M. (2009b). Stimulus exposure and gaze bias: A further test of the gaze cascade model. *Attention, Perception and Psychophysics*, 71 (3), 445-450.

Glaholt, M. G., & Reingold, E. M. (2011). Eye movement monitoring as a process tracing methodology in decision making research. *Journal of Neuroscience, Psychology, and Economics*, 4(2), 125-146.

Glaholt, M. G., Wu, M.-C., & Reingold, E. M. (2010). Evidence for top-down control of eye movements during visual decision making. *Journal of Vision*, 10 (5), 1-10.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13 (10), 1292-1298.

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, 108 (33), 12852-13857.

Liao, H.-I., & Shimojo, S. (2012). "Dynamic preference formation via gaze and memory," in R. J. Dolan and T. Sharot (Eds.), *Neuroscience of Preference and Choice*. San Diego: Elsevier Academic Press. pp. 277-292.

Liao, H.-I., Yeh, S.-L., & Shimojo, S. (2011). Novelty vs. familiarity principles in preference decisions: task-context of past experience matters. *Frontiers in Psychology*, 2, 1-8.

Meloy, M. G., & Russo, J. E. (2004). Binary choice under instructions to select versus reject. *Organizational Behavior and Human Decision Processes*, 93, 114-128.

Nittono, H., & Wada, Y. (2009). Gaze shifts do not affect preference judgements of graphic patterns. *Perceptual and Motor Skills*, 109, 79-94.

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

Park, J., Shimojo, E., & Shimojo, S. (2010). Roles of familiarity and novelty in visual preference judgments are segregated across object categories. *Proceedings of the National Academy of Sciences*, 107 (33), 14552-14555.

Reidl, R., Brandstätter, E., & Roithmayr, F. (2008). Identifying decision strategies: A process- and outcome-based classification method. *Behavior Research Methods*, 40 (3), 795-805.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, 108 (2), 370-392.

Russo, J. E. (2011). "Eye fixation as a process trace" in M. Schulte-Mechklenbeck, A. Kuhberger, & R. Raynard (Eds.), *A Handbook of Process Tracing Methods for Decision Research*. New York: Psychology Press, pp. 43-64.

Schotter, E. R., Berry, R. W., McKenzie, C. R. M., & Rayner, K. (2010). Gaze bias: Selective encoding and liking effects. *Visual Cognition*, 18 (8), 1113-1132.

Schotter, E. R., Gerety, C., & Rayner, K. (2012). Heuristics and criterion setting during selective encoding in visual decision making: Evidence from eye movements. *Visual Cognition*, 20 (9), 1110-1129.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, 6 (12), 1317-1322.

Simion, C., & Shimojo, S. (2006). Early interactions between orienting, visual sampling and decision making in facial preference. *Vision Research*, 46 (20), 3331-3335.

Simion, C., & Shimojo, S. (2007). Interrupting the cascade: Orienting contributes to decision making even in the absence of visual stimulation. *Attention, Perception and Psychophysics*, 69 (4), 591-595.

Tsetsos, K., Usher, M., & Chater, N. (2010). Preference reversal in multiattribute choice. *Psychological Review*, 117 (4), 1275-1293.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108 (3), 550-592.

Figure captions.

Figure 1. Examples of the visual stimuli (a) faces (b) scenes

Figure 2. Gaze biases in (a) total dwell duration, (b) number of dwells, (c) mean dwell duration and (d) as a proportion of total dwell duration. Error bars represent standard error of the mean.

Figure 3. Task and stimulus effect in Gaze likelihood curve (a) Face like task v.s. Face dislike task, (b) Scene like task v.s. Scene dislike task, (c) Face like task v.s. Scene like task and (d) Face dislike task v.s. Scene dislike task. Dotted lines represent 95% confidence intervals about the mean.

Figure 4. Probability of looking the chosen item last in the trial. Error bars represent standard error of the mean.

Figure 5. Dwell sequence analysis for (a) Face like task, (b) Scene like task, (c) Face dislike task and (d) Scene dislike task. Error bars represent standard error of the mean.

Figure 6. (a) First encounter duration, (b) Proportion of trials separated by the total number of dwells in the trial, (c)to(f): First encounter duration (c) for trials with two dwells, (d) trials with three dwells, (e) trials with four dwells, and (f) trials with five or more dwells

Figure 7. Dwell duration from the beginning in the trial in which (a) the first dwell was

directed to the chosen item, (b) the first dwell was directed to the other item. Each line shows the change of dwell duration for trials with different number of dwells. The legend show the change of the dwelled item (c: chosen item, a: the other item). For example, c-a-c shows the dwell duration for trials in which the first dwell was directed to the chosen item, the next dwell was directed to the other item and the last dwell was directed to the chosen item. Error bars represent standard error of the mean.