

Task-Dependent Eye-Movement Patterns in Viewing Art

Nino Sharvashidze Alexander C. Schütz
Philipps-Universität Marburg, Philipps-Universität Marburg,
Germany Germany

In art schools and classes for art history students are trained to pay attention to different aspects of an artwork, such as art movement characteristics and painting techniques. Experts are better at processing style and visual features of an artwork than nonprofessionals. Here we tested the hypothesis that experts in art use different, task-dependent viewing strategies than nonprofessionals when analyzing a piece of art. We compared a group of art history students with a group of students with no art education background, while viewing 36 paintings under three discrimination tasks. Participants were asked to determine the art movement, the date and the medium of the paintings. We analyzed behavioral and eye-movement data of 27 participants. Our observers adjusted their viewing strategies according to the task, resulting in longer fixation durations and shorter saccade amplitudes for the medium detection task. We found higher task accuracy and subjective confidence, less congruence and higher dispersion in fixation locations in experts. Expertise also influenced saccade metrics, biasing it towards larger saccade amplitudes, advocating a more holistic scanning strategy of experts in all three tasks.


Keywords: art perception, task-dependent viewing, expertise, confidence, viewing strategies, eye movements, scanpath, eye tracking

Introduction

Task-driven selection

It is widely agreed that visual attention is on one hand stimulus-driven, induced by the physical properties of an image, referred to as salience, categorized as bottom-up processing. On the other hand, top-down demands like current goals, task, knowledge, expectations and reward also direct our attention (for review see Schütz et al., 2011; Tatler et al., 2011). Stimulus-driven models of attention claim

that where the observer first looks at in a scene is determined by the low-level features of a scene. Strong contrasts in brightness, color and orientation stand out from the rest of the scene and grab the attention of the viewer (for review see Itti & Koch, 2001). Bottom-up processing has been shown to drive our attention due to low level feature correlation with object properties (Einhäuser et al., 2008b). Generally, bottom-up models have been more successful in explaining fixations in free viewing (e.g. Parkhurst et al., 2002) but they come to their limits when observers start to pursue specific goals (e.g. Einhäuser et al., 2008a). Previous research has shown that salience can explain only a small portion of fixations in everyday tasks. During everyday activities like driving (Land & Lee, 1994), making tea (Land et al., 1999) or sandwiches (Hayhoe, 2000) observers only fixate task-relevant areas and ignore task-irrelevant, although salient areas. Buswell (1935) was probably first to notice that, “The mental set obtained by the directions given for looking at a picture

Received July 30, 2020; Published December 16, 2020.
Citation: Sharvashidze, N. & Schütz, A. C. (2020). Task-Dependent Eye-Movement Patterns in Viewing Art. *Journal of Eye Movement Research*, 13(2):12.
Digital Object Identifier: 10.16910/jemr.13.2.12
ISSN: 1995-8692
This article is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). 

obviously influences the characteristics of the perceptual process” (Buswell, 1935, p.136). However, one of the most famous examples of task-driven visual guidance comes from Yarbus (1967) who used paintings to demonstrate the influence of different tasks on the eye movements. For instance, if the observer was asked to remember the clothes of depicted people, mainly clothes were attended, whereas if the observer had to estimate the age of people, the most informative areas for this particular task—faces, were fixated. His demonstration has been replicated a few times under more valid conditions with some modifications leading to comparable results (Borji & Itti, 2014; DeAngelus & Pelz, 2009; Tatler et al., 2010).

Fixation count and fixation durations, saccade count and saccade amplitudes are useful estimates for the task influence on scene processing (Cronin et al., 2020; Mills et al., 2011; Nuthmann, 2016; Nuthmann et al., 2010; Võ & Henderson, 2009). Castelhana et al. (2009) found that task affected where the observers fixated. Fixations were widely distributed in the memorization task and more focused on search-relevant areas in the search task. They were also able to find differences in the saccade amplitude during the initial viewing of the scene. Dynamics of fixation and spatial fixation density seem to be important factors that are variable as a function of a task (for a recent review see Williams & Castelhana, 2019).

Numerous different tasks have been used to explore task-dependent visual selection while viewing art, including free viewing (Koide et al., 2015; Pang et al., 2013), judging the aesthetic quality (Fudali-Czyż et al., 2018; Massaro et al., 2012; Molnar, 1981; Pihko et al., 2011), paying attention to semantic content (Molnar, 1981), memorization (Zangemeister et al., 1995), categorization and person detection (Wallraven et al., 2008), movement detection (Massaro et al., 2012). Despite the diversity of tasks, to the best of our knowledge no common art historian tasks have yet been used to examine task-driven viewing patterns while looking at paintings. Conventional art expertise tasks, including accessing the value or the authenticity of the artwork go hand in hand with the recognition of style and art movement, medium and technique, as well as the correct dating of the artworks. Accurate assessment of the art movement requires knowledge about art periods, as well as understanding of characteristic categories of depicted subjects and content and distinctive modes of expression. Date estimation involves some further back-

ground information about historical context and the timeline of art history. Defining the physical characteristics of the painting includes determining the materials and tools that artists use to create a work of art. The knowledge required to solve all these tasks are passed on in art schools and classes for art history.

Expertise

Expertise is broadly defined as constantly superior performance within a specific domain relative to novices (Ericsson & Lehmann, 1996). Expertise criteria used in different studies in the visual domain include performance efficiency, social recognition, group membership and years of experience (Gegenfurtner et al., 2011). By trying to summarize what contributes to the expert superior visual performance, Gegenfurtner et al. (2011) considered three theories supported by the literature in their meta-analysis (see also Brams et al., 2019). The theory of long-term working memory (Haider & Frensch, 1999) assumes that experts encode and retrieve information more rapidly than novices. According to this theory, experts’ rapid information processing should be reflected in shorter fixation durations. The information-reduction hypothesis (Ericsson & Kintsch, 1995) focuses on learned selectivity of information processing. It proposes that expertise optimizes the amount of processed information by neglecting task-irrelevant information and focusing on task-relevant information. This results in more fixations on task-relevant areas and fewer on irrelevant areas. The holistic model of image perception (Kundel et al., 2007), proposes the idea that experts are able to extract information from widely distanced and parafoveal regions. This model anticipates larger saccade amplitudes and shorter time to first fixate task-relevant areas in experts.

Where/what does one have to attend in order to extract useful information seems to depend on the scene context and knowledge of the scene statistics. Scene statistics that are learned from experience are linked to full range of selection history effects (Awh et al., 2012; Chun & Jiang, 1998; Theeuwes, 2018; Theeuwes, 2019). Experts’ ability to extract task-relevant information is linked to their knowledge of scene context and scene statistics based on past visual experience and declarative knowledge about the paintings. Generic semantic and spatial knowledge about the particular type of scene is described as scene-schema knowledge (Henderson, 2003). For example, it includes information about the objects that are likely to be found in this scene category. Scene-schema can be rapidly

retrieved and used to limit fixations to scene areas that are likely to contain task-relevant information. This could be information about how paintings are generally organized, about the typical compositional structures and the common subjects that are depicted. Another type of knowledge, task-related knowledge can involve a general “gaze-control policy” or strategy relevant to a given task (Henderson, 2003), knowledge about style specific regularities, understanding of how the mixes of pigments and binding media usually look like, how they are applied and where they can be conveniently identified. The same assumption is illustrated in the Leder et al.’s (2004) model of aesthetic experience, where past experience with art, for example, exposure to artworks and declarative knowledge can impact the experience on separate levels. On the one hand, previous experience, which also interacts with domain specific expertise affects implicit memory integration, on the other hand, domain specific declarative knowledge influences explicit classification and enables, content vs. style processing. Therefore, in the case of art experts, one encounters the combination of visual experience, due to the exposure to many paintings throughout art classes and a comprehensive declarative knowledge about art.

Augustin and Leder (2006) showed that, while non-experts merely use terms familiar to them from everyday experience, experts relate artworks according to the style, art movement or other art-specific aspects. Buswell (1935), who was the first to compare fixation patterns of artistically trained and untrained subjects, could not find any systematic differences between the groups. He was, however, quite tentative about the nonexistence of the differences and remained critical of his own methods. Indeed, later research has provided some evidence that experts and non-experts differ in how they look at paintings. Nodine et al. (1993) analyzed eye movements of professional and nonprofessional viewers, who judged compositions that differed in balance through manipulation of composition symmetry. They found a more global strategy of scanning an image in experts. The same pattern was found by Zangemeister et al. (1995). Experts also tend to make less frequent and shorter fixations on narrative elements (Kristjanson & Antes, 1989; Pihko et al., 2011; Vogt & Magnussen, 2007). Ylitalo et al. (2016) found that experts tend to have shorter fixation durations and less variability in fixation durations than novices. Fedorovskaya et al. (2017) found less local viewing by experts and less re-fixations on already attended areas. Pihko et al. (2011) demonstrated that gaze patterns of non-experts became closer to those of

experts after they received additional information about the painting.

To sum up, evidently, diverse tasks have been investigated in the context of visual art, however, previous studies have mainly focused on free viewing and aesthetic judgment tasks. Art-specific tasks have been overlooked. Art historians, experts in conservation and restoration encounter domain-specific tasks on a daily basis. However, we do not know how the specific task demands affect their eye movement patterns and viewing strategies. Eye movement parameters and spatial fixation density seem to vary from task to task, thus we assume that art-specific tasks will also induce variation in fixation dynamics and attended painting areas. Regarding expertise, we assume that experts, having scene-schema knowledge, task-specific and declarative knowledge about the paintings, establish different task-dependent strategies than non-experts. It is likely, that they generally, develop hypotheses regarding a painting and control gaze in order to efficiently solve the problem. For instance, Nodine et al. (1993) showed that experts are sensitive to composition and use structure of a painting to allocate their attention to the informative parts of the painting. Bauer and Schwan (2018) showed that experts are more effective in searching for content that is helpful for successful meaning-making in the Renaissance portraits. Accordingly, experts seem to be better at looking for specific aspects of an artwork that are relevant for the task.

In this study, we examined experts and non-experts in art by giving them three different art related tasks to solve. Participants were asked to define the style/art movement, estimate the date and discriminate the medium of the paintings. In other words, they were asked to identify how, when and whereby the painting was created. The major disadvantage of stimulus sets used to date is the inclusion of paintings with highly salient areas, including faces and texts. With this in mind, we aimed to create a better-suited set of stimuli omitting highly salient face areas and incorporated texts, which are known to quickly attract gaze (e.g. Cerf et al., 2009) and can, in some cases, work as cues for the tasks. We aimed to explore familiarity of paintings, accuracy and confidence ratings of our observers. We expected to see task-driven effects on spatial and temporal characteristics of eye movement behavior.

Methods

Participants

In total 29 subjects (age: $M = 25.06$, $SD = 3.8$, range: 19–35 years, 5 male, 3 left-handed) participated in our study. 13 Experts (age: $M = 25.38$, $SD = 3.25$, 2 male) were recruited from art related Marburg University courses including History of Art B.A. (5), History of Art M.A. (6), Concepts of Fine Arts M.A. (1) and Visual Arts, Music and Modern Media B.A. (1). 16 non-experts (age: $M = 24.81$, $SD = 4.33$, 3 male) were students of the following art unrelated university courses: Psychology (6), Political Science (2), Business Administration (1), Sociology (1), Human Medicine (1), German Language and Literature (1), Islamic Studies (1), Educational Science (1), Romance Studies (1) and German as a Foreign Language (1). The data from two non-expert subjects had to be excluded. One due to the technical issue during the experiment (Subject 20), the other due to the poor quality of the eye movement data (Subject 33), resulting in a total sample size of 27 subjects with 13 experts and 14 non-experts. The aforementioned art related university courses offer “Basics of Art History” (History of Art B.A.) or “Propaedeutics of Art” (Visual Art, Music and Modern Media B.A.) as a compulsory class in the first semester. Over the whole course more classes are offered, providing the graduates with thorough knowledge of the important artworks, genres and techniques from late antiquity to the present. All subjects had normal or corrected to normal vision and gave informed consent prior to participation. All of them had normal color vision according to the Ishihara Test (Ishihara, 1960). They were paid 8€/h and received a book at the end of the experiment. The experiment was in accordance with the principles of the 1964 Declaration of Helsinki and was approved by the ethics committee of the Marburg University, Department of Psychology (proposal 2017-27k).

Apparatus

The experiment was conducted using the Psychtoolbox (Brainard, 1997; Pelli, 1997) in MATLAB (R2016a; The MathWorks, Natick, MA) and stimuli were presented on a VIEWPixx monitor (VPixx Technologies, Inc, Saint-Bruno, Quebec, Canada) at a viewing distance of 60 cm. The monitor had a spatial resolution of 1920 x 1080 pixel and a size of 51.5 x 29 cm. Eye movements of the right eye were recorded at 1000 Hz using a desktop mounted EyeLink 1000 (SR Research Ltd., Ontario, Canada) and the

EyeLink Toolbox (Cornelissen et al., 2002). Due to technical issues, the data of five participants were recorded at 2000 Hz and downsampled to 1000 Hz before further analysis.

Materials

The stimuli of the experiment consisted of 36 digitized images of paintings taken from 12 different databases (Art Institute Chicago, Barnes Foundation, Brooklyn Museum, Lithuanian Art Museum, Mauritshuis, Metmuseum, Philadelphia Museum of Art, Rijksmuseum, The Athenaeum, Wikiart, Wikimedia Commons and Yale Center for British Art). The Athenaeum database is no longer available online. In order to prevent solving tasks based on merely image background information we chose lesser-known paintings with no or illegible signatures. We avoided paintings of human figures with highly salient face areas and incorporated texts. Images had a maximum resolution of 1300 x 800 pixels and were presented in front of a grey background. In order to sufficiently balance the stimuli, we narrowed down the art movement range of selected paintings to six art movements: Baroque, Expressionism, Impressionism, Cubism, Post-Impressionism, Romanticism and three media used: Oil, Pastel and Watercolor. Our stimuli set consisted of six paintings per six art movements. 12 of these paintings were oil paintings, 12—watercolor paintings and 12—pastel paintings. The first requirement for the successful stimuli-task assignment to each subject was that no participant could view the same painting twice. This way we could eliminate memory biases masking the task-depending viewing. The second requirement was that all paintings had to be shown under all three task conditions to ensure no differences occurred from image properties only. To fulfil these requirements, we first divided our 36 paintings in three sets with balanced style and medium categories. For each participant, each set was assigned with a different task. Images mapped to a certain task were then presented randomly in the experiment. Since we only had three sets and tasks, six participants would exhaust all, in this case, possible task-set combinations. To avoid the replication of task-set configurations a second arrangement of images to sets was introduced. We kept the balance of styles and media, but shuffled the positions of the paintings in the second arrangement. In the end, 12 participants per group could ensure the fulfilment of both requirements with 36 paintings rearranged in two different ways (see supplementary online

material for more details). Information about date and medium of paintings was available from databases. Further, some databases (e.g. Wikiart) provided information about the style. We examined whether the provided information was credible by tracing the history timeline and representative artistic style information. Most of the time, given the date, technique used and subject matter of the painting, it was possible to define an art movement style. We used standard art movement definitions available in art history books (e.g. Gombrich, 1966).

To make sure that our operationalization of expertise was successful, we used an art expertise questionnaire from Pang et al. (2013). Despite the lack of knowledge-related questions, according to Pang et al. (2013), this measure covers the formal art education part of the expertise in greater detail compared to similar measures. It also contains questions about the amount of time spent on interaction with art and questions about formal art analysis skills. We analyzed 19 visual art-related questions with numerical information, the answer scores were summed up.

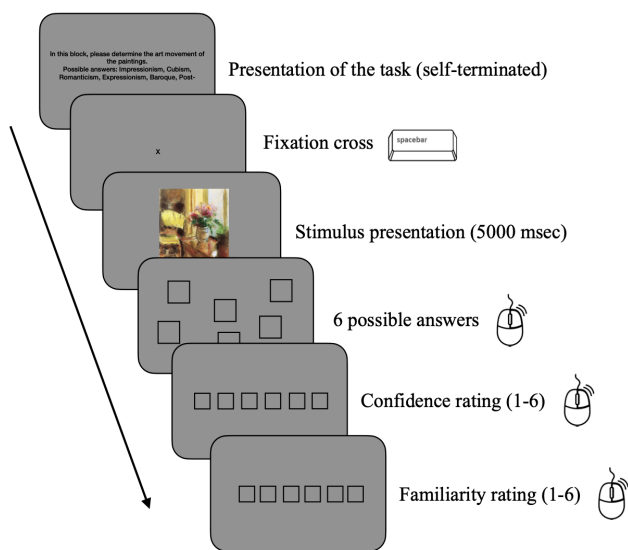


Figure 1: Trial Procedure. Each block started with a self-terminated task presentation followed by trials starting with a fixation cross. Trials consisted of a self-terminated fixation on a cross in the center of the screen, followed by the stimulus presentation for 5000 ms. After stimulus offset, the desired answer could be chosen from six spatially randomly presented answers with a mouse click. Subsequently confidence and familiarity ratings appeared successively on the screen and were entered using the mouse.

Procedure

Participants were asked to fill in the consent form and answer the demographic questions. They were instructed about the experiment and tested for color deficiencies. The experiment consisted of three blocks, each containing 12 trials. The eye tracker was calibrated at the beginning of each block. Each block began with the presentation of the task and six possible answers. By fixating the fixation cross in the center of the screen and pressing the space bar participants started the presentation of the stimulus. Each painting was presented for 5000 milliseconds. The presentation could be terminated earlier by pressing the space bar, but only six participants (three experts, three non-experts) skipped the 5000 milliseconds viewing time. Participants could choose one correct answer from six possible answers with a mouse click. The six possible answers for the art movement task were Baroque, Expressionism, Impressionism, Cubism, Post-Impressionism, Romanticism. The six possible answers for the medium task were Oil, Pastel, Watercolor, Acrylic, Chalk and Ink. The six possible answers for the date task were chosen to roughly match the corresponding art periods: 1600-1730, 1904-1921, 1860-1910, 1912-1924, 1889-1935, 1762-1836. The answers were randomly presented on the screen to eliminate the position biases. Participants were also asked how confident they were in their answer on a scale of 1 to 6 (1-very confident, 6-not confident) and how familiar was the painting on a scale of 1 to 6 (1-very familiar, 6-unfamiliar), ratings of familiarity and confidence were later reversed for the analysis (1-not confident, 6-very confident and 1-unfamiliar, 6-very familiar), so that the interpretation could be more intuitive. Participants could take a small break if needed before the calibration for the next block started. After the experiment they were asked to fill in the questionnaire.

Design

The following independent variables were considered for the analysis: 1) expertise: experts (students of art related university courses) and non-experts in art (students of art unrelated university courses including social sciences and humanities), 2) task (style/art movement, date, medium). As dependent variables we considered: 1) accuracy of the task response, 2) familiarity rating of the painting, 3) confidence rating about the response, 4) fixation duration (ms), 5) saccade amplitude (deg).

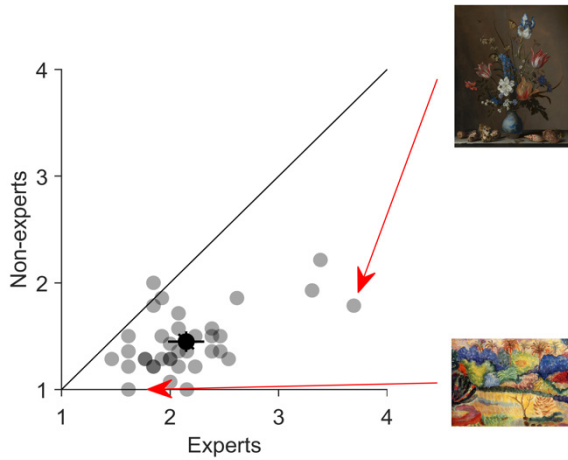


Figure 2: Average familiarity ratings of experts and non-experts for 36 paintings. Half-transparent symbols indicate single images, the non-transparent symbol the mean across images. Error bars correspond to 95% confidence intervals. The paintings with the highest absolute familiarity value (painting by Balthasar van der Ast (29)) and with the lowest absolute familiarity value (painting by Paul Gauguin (6)) are shown on the right.

Analyses

We used the EyeLink 1000 algorithm with a combined velocity threshold of $30^\circ/\text{sec}$ and acceleration threshold of $8000^\circ/\text{sec}^2$ to determine saccades. Fixations were detected if they fulfilled the criterion of not belonging to a saccade. We removed blinks and fixations outside the image. Fixations immediately preceding and following blinks were discarded. 120 saccades with unusual trajectories, long durations given their amplitudes ($\text{duration}/\text{amplitude} > 1.5$) or notably large distances traveled by the eye given their amplitudes ($\text{distance}/\text{amplitude} > 3.5$) were also excluded from the analysis. We used R (R Core Team, 2020) and lme4 package (Bates et al., 2015) to perform a linear mixed effects analysis (LME). We estimated four models with the dependent variables: accuracy, confidence, fixation duration and saccade amplitude. We entered the independent variables task and expertise as fixed effects, without interaction term into the models. As random effects, we entered intercepts for subjects and images. P-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question. The R package rmcrr (Bakdash & Marusich, 2017) was used to calculate repeated measures correlation between

Table 1: Mean values (M) and standard deviations (SD) of dependent variables, Normalized Scanpath Saliency (NSS) and Dispersion for both groups

	Experts		Non-experts	
	M	SD	M	SD
Expertise score (max. 70)	39.00	5.00	21.19	6.30
Accuracy	0.50	0.50	0.36	0.48
Confidence	4.25	1.36	3.34	1.48
Familiarity	2.15	1.45	1.45	1.02
Fixation duration (ms)	267.36	126.80	265.83	124.40
Saccade amplitude ($^\circ$)	5.55	1.40	5.05	1.26
NSS	1.17	0.44	1.40	0.44
Dispersion	0.15	0.04	0.14	0.04

familiarity and confidence. It accounts for non-independence by statistical adjustment of inter-individual variability.

Data visualization

In order to provide an overview about the spatial distribution of eye movement parameters, we first visualized the scanpath per trial by plotting fixations and saccades on the corresponding paintings. To illustrate the fixation data, we aggregated fixation density maps, so-called heatmaps, for experts and non-experts, where all fixations for a single image divided by the sum of fixation durations were merged to form a density map. This fixation maps were smoothed by convolving a Gaussian kernel. They show the accumulated time observers spent looking at the certain areas of the paintings by representing the values as colors. For 36 images, fixation densities across 13 experts and 14 non-experts for three tasks were aggregated into six maps.

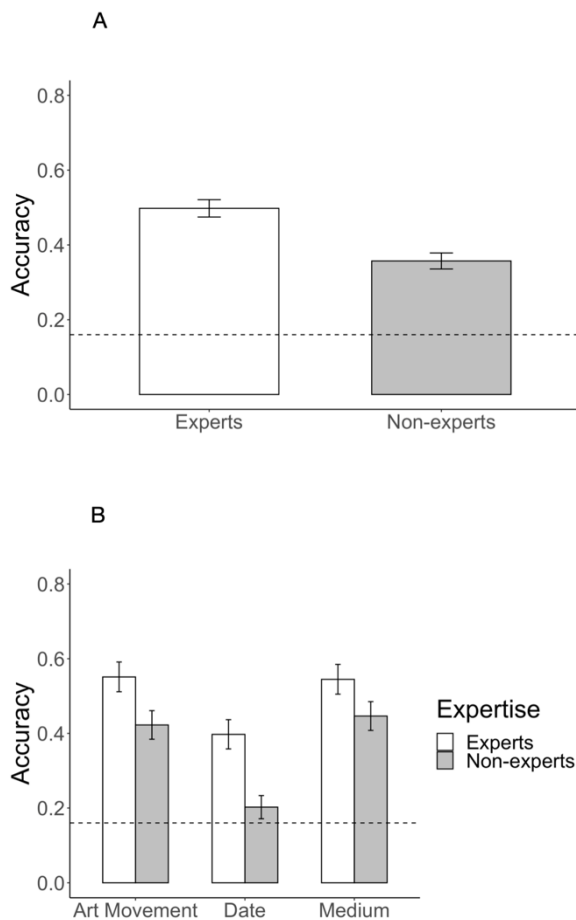


Figure 3: (A) Mean accuracy of experts and non-experts, (B) mean accuracy of experts and non-experts across all three tasks. Dashed line shows chance level of performance. Error bars correspond to 95% confidence intervals.

In order to facilitate the comparison additionally three difference density plots for both groups for each image were aggregated.

Normalized Scanpath Saliency and Dispersion

Heatmaps showed that fixation distributions for the task medium had a slightly different appearance. After the detection of visually observable differences, we took a closer look and used the Normalized Scanpath Saliency (NSS) method, initially introduced by Peters et al. (2005), to calculate the similarity of fixations across observers. Analogous to Dorr et al. (2010), who used NSS method to

evaluate the similarity of eye movements of multiple observers, we used a standard machine learning method “leave one out”. For each observer with a given task on a given image, the fixations of all other observers with the same task on the same image were used to create a fixation map. To this end, each fixation was modelled as a Gaussian with a standard deviation of 1.5° and all Gaussians were summed. This map was then normalized to have a mean of zero and a variance of unity. Finally, the fixation values at the fixations of the left-out observer were summed. Positive NSS indicates congruence between the regions fixated by one observer and all other observers. Zero values indicate uncorrelated fixations and negative NSS points out incoherent fixation locations. We also calculated dispersion – the size of spatial fixation distribution for experts and non-experts across all three tasks for all images as described in Holmqvist and Andersson (2017). Each fixation was rated using a Gaussian with a standard deviation of 1.5° . For each trial all Gaussians were summed up and normalized to a maximum of 1. The values were then divided by the number of pixels so that the values could reside between 0 and 1.

Results

Behavioral parameters

Experts achieved higher art expertise scores ($M = 39.00, SD = 5.00$) than non-experts ($M = 22.00, SD = 6.30$). The group difference was highly significant ($t(24.49) = -7.82, p < 0.001$). According to this result, our participant allocation to both groups might be regarded as reliable. We also found differences in the average task accuracy between two groups (Fig. 3, Table 1). Experts were more accurate in their answers than non-experts ($\chi^2(1) = 9.31, p < 0.01$). As chance performance would be at 0.16, both expert and non-expert performances were significantly above chance. Accuracy also varied between tasks ($\chi^2(2) = 38.67, p < 0.001$). Participants of this study had most difficulties with dating the paintings correctly ($M = 0.30, SD = 0.46$) and were comparably accurate in the tasks – art movement ($M = 0.48, SD = 0.50$) and medium ($M = 0.49, SD = 0.50$). These differences in accuracy between groups and tasks were also reflected in the subjective confidence ratings. Reported confidence (Table 1) was lower by about 0.90 in non-experts compared to experts ($\chi^2(1) = 8.51, p < 0.01$). Subjective confidence varied between tasks ($\chi^2(2) = 71.31, p < 0.001$). Highest confidence was reported on the

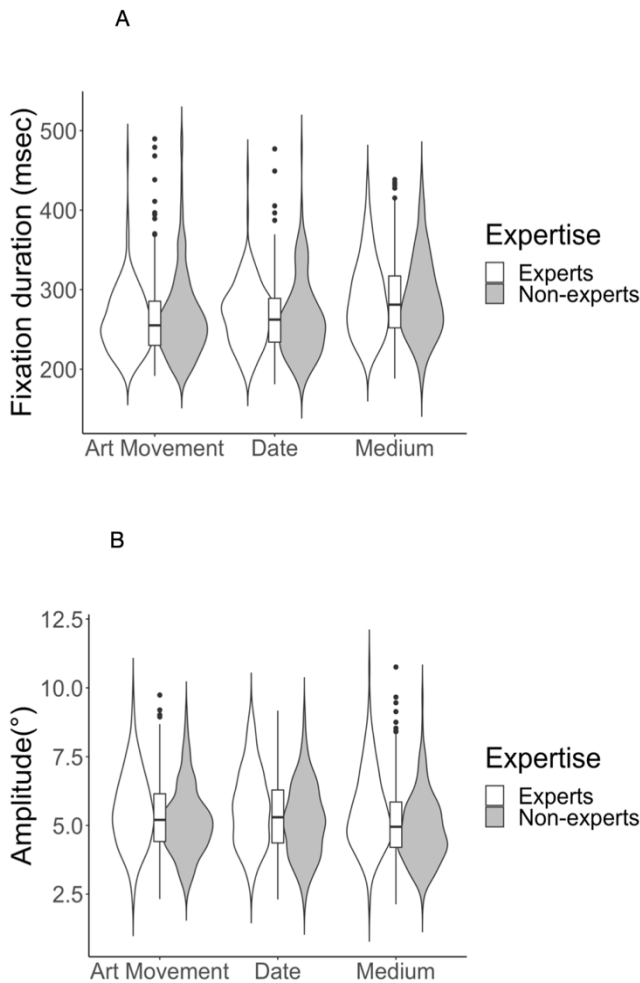


Figure 4: (A) Average fixation duration and (B) average saccade amplitude for experts and non-experts across three tasks.

medium task ($M = 4.19$, $SD = 1.46$), followed by art movement ($M = 3.71$, $SD = 1.49$) and date tasks ($M = 3.43$, $SD = 1.44$). Reported confidence for correct answers ($M = 4.31$, $SD = 1.40$) was higher, than reported confidence for incorrect answers ($M = 3.39$, $SD = 1.43$). A correlational analysis revealed a weak correlation between confidence and familiarity ($r(944) = 0.21$, $p < 0.001$), indicating that subjects reported higher confidence on the more familiar paintings and lower confidence on the less familiar paintings. Overall, the reported familiarity of the paintings was quite low across all participants ($M = 1.78$, $SD = 1.29$). Our selection of paintings was even less familiar for non-experts than experts ($t(35) = 10.105$, $p < 0.001$). These results

suggest the adequate selection of the stimuli in our experiment (Fig. 2).

Eye movement parameters

Fixations lasted on average 266.56 ms ($SD = 125.55$). The amplitude of saccades was on average 5.29° ($SD = 3.80$). These results are in line with previous work reporting an average saccade amplitude of 4° – 6° on scenes that are on average 20° – 30° wide (e.g. Castelano et al., 2009; Unema et al., 2005) and average fixation duration of 200–300 ms (e.g. Rayner, 2009) in scene perception tasks. We could not find any effect of expertise on fixation duration (Table 1). In order to normalize the distribution, we subjected fixation duration to logarithmic transformation for the LME analysis. Whereas expertise did not have an effect on fixation duration, task influenced fixation duration ($\chi^2(2) = 122.21$, $p < 0.001$). The highest average fixation duration was detected during the medium task ($M = 282.38$, $SD = 129.88$), followed by the date task ($M = 260.76$, $SD = 123.76$) and the art movement task ($M = 257.37$, $SD = 121.70$). These results suggest that participants' fixations lasted longest when their task was to detect what medium was used for the painting. We found no significant interaction between expertise and task for fixation duration. Task also influenced saccade amplitude ($\chi^2(2) = 13.87$, $p < 0.001$) increasing it by about 0.05° for task date ($M = 5.39$, $SD = 1.37$) compared to art movement task ($M = 5.34$, $SD = 1.31$) and reducing it by about 0.19° for task medium ($M = 5.15$, $SD = 1.37$) compared to art movement task. Consequently, in addition to longer fixation durations, task medium also induced smaller saccade amplitudes. This means that on average, our participants fixated longer and their eyes traveled a shorter distance during the medium task. We also found that in the medium task saccade amplitudes of non-experts were smaller ($M = 4.83^\circ$, $SD = 1.22$), than experts' ($M = 5.49^\circ$, $SD = 1.44$), (Table 2). However, no significant interaction between expertise and task could be found using the LME model for saccade amplitude. We could also detect some differences between average saccade amplitudes on the group level (Table 1). Expertise had an effect on saccade amplitude ($\chi^2(1) = 4.67$, $p < 0.05$) reducing it by about 0.50° in non-experts compared to experts. From these results we can assume that expert group explored the paintings more extensively, as their eyes traveled larger distances across all three tasks.

Normalized Scanpath Saliency and Dispersion

Variability in eye movement patterns was measured using the Normalized Scanpath Saliency method. We found rather high NSS values in our data ($max = 2.69, M = 1.28$) indicating an overall high coherence in fixation patterns between observers. Averaged NSS values for both groups are shown in Table 1, averaged NSS values for three tasks split by expertise are shown in Table 2. The LME model with expertise as fixed effect and image as random intercept revealed significantly higher average expert NSS score compared to average non-expert NSS score ($\chi^2(1) = 43.89, p < 0.001$) indicating larger variability of eye fixation distribution in experts. We also found the smallest mean NSS value for the task medium ($M = 0.97, SD = 0.35$) compared to two other tasks. The LME model with task as fixed effect and image as random effect suggested less coherence in eye movement patterns for the task medium ($\chi^2(2) = 131.35, p < 0.001$), (Fig 5). The LME model with expertise as fixed effect and image as random intercept revealed higher dispersion values ($\chi^2(1) = 21.42, p < 0.001$) by experts ($M = 0.15, SD = 0.04$) compared to non-experts ($M = 0.14, SD = 0.04$), indicating that they fixated larger areas of the paintings, compared to non-experts. The LME model with task as fixed effect and image as random

intercept revealed that fixation dispersion significantly varied between tasks ($\chi^2(2) = 43.47, p < 0.001$). The highest spatial fixation dispersion value was found for the medium detection task ($M = 0.16, SD = 0.04$), compared to art movement ($M = 0.14, SD = 0.04$) and date ($M = 0.14, SD = 0.04$) tasks, suggesting that both experts and non-experts fixated larger areas of the paintings during the medium detection task. The dependent values for all three tasks split by expertise can be found in Table 2.

Temporal differences

To investigate temporal information, we plotted fixation durations and saccade amplitudes over the course of a trial. Previous research has shown that the first fixation contains a strong central bias (Tatler, 2007) and differs from later fixations (e.g. Over et al., 2007). The average duration of the initial fixation in our data ($M = 384.97, SD = 130.63$) was longer than the overall average fixation duration ($M = 266.56, SD = 125.55$). The average amplitude of the initial saccade ($M = 3.19^\circ, SD = 1.89$) was smaller than the average saccade amplitude ($M = 5.29^\circ, SD = 1.35$). Similar results were reported by Over et al. (2007) and Van Loon et al. (2002). We therefore excluded the first fixations and first saccades from the analysis and investigated fixations and saccades in the ordinal number range

Table 2: Mean values (M) and standard deviations (SD) of dependent variable, Normalized Scanpath Saliency (NSS) and Dispersion for three tasks split by Expertise

	Art movement				Date				Medium			
	Experts		Non-experts		Experts		Non-experts		Experts		Non-experts	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Accuracy	0.55	0.50	0.42	0.50	0.40	0.49	0.20	0.40	0.55	0.50	0.45	0.50
Confidence	4.09	1.45	3.36	1.44	4.04	1.23	2.86	1.38	4.62	1.33	3.80	1.47
Fixation duration (ms)	260.83	40.47	265.42	52.14	269.26	39.92	263.49	48.97	288.79	48.61	288.50	50.49
Saccade amplitude ($^\circ$)	5.55	1.36	5.15	1.24	5.62	1.39	5.17	1.31	5.49	1.44	4.83	1.22
NSS	1.37	0.42	1.58	0.39	1.28	0.42	1.54	0.40	0.86	0.31	1.08	0.35
Dispersion	0.15	0.04	0.13	0.03	0.15	0.04	0.14	0.04	0.16	0.04	0.15	0.04

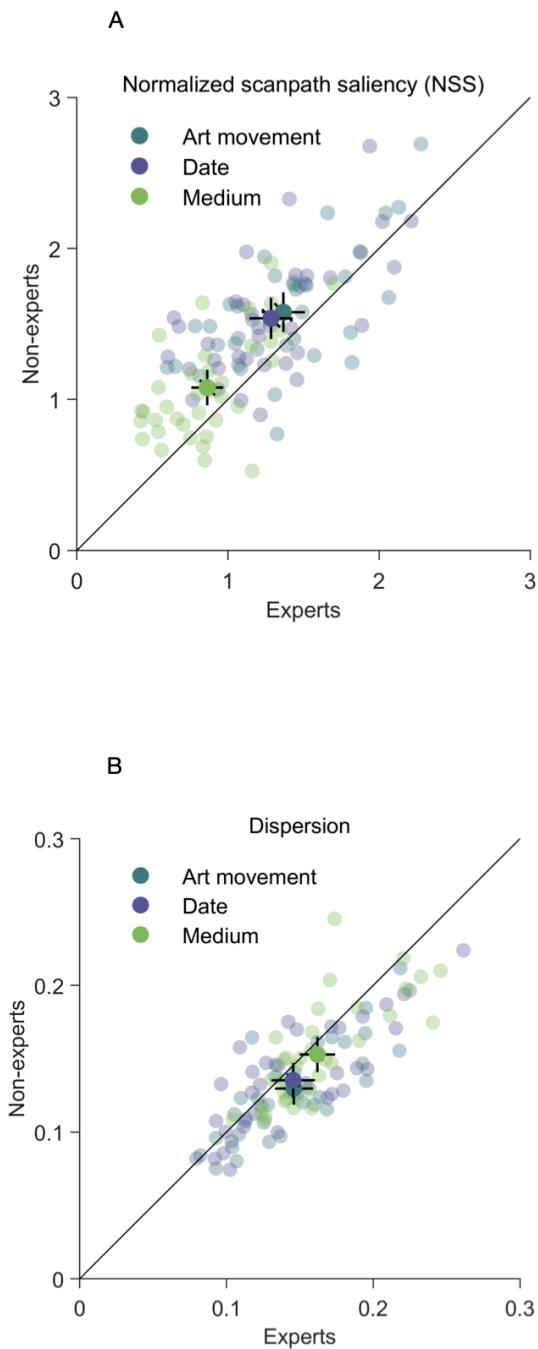


Figure 5: Average NSS (A) and Dispersion (B) values, within groups, within tasks comparison. Half-transparent symbols indicate single images, the non-transparent symbols the mean across images. Error bars correspond to 95% confidence intervals.

of 2–13. As expected, we found increasing fixation durations as a function of ordinal fixation number. The LME model with fixation number as fixed effect and subject and image as random effects revealed a significant effect of ordinal fixation number on fixation duration ($\chi^2(11) = 370.37, p < 0.001$).

We also observed amplitudes becoming shorter across the trial. The LME model with saccade number as fixed effect and subject and image as random effects revealed a significant effect of ordinal saccade number on saccade amplitude ($\chi^2(11) = 69.00, p < 0.001$). Fixation durations and saccade amplitudes for each task plotted by ordinal fixation and ordinal saccade number and fixation durations and saccade amplitudes for experts and non-experts plotted by ordinal fixation and saccade number are shown in Figure 6.

Discussion

In this study, we examined the eye movements of observers with art education and observers without art education background, while they were viewing 36 digitized paintings under three task conditions. Participants were asked to identify the style/art movement, date or medium of the paintings. We first looked at the behavioral parameters. The questionnaire results reassured the optimal allocation of participants to the groups (Table 1). Accuracy test showed that experts, although not being highly accurate, outperformed non-experts and chance level of accuracy (Fig. 3). Experts' superior performance in style/art movement detection supports Augustin and Leder's (2006) finding that experts are better at style processing and is in line with Leder et al.'s (2004) model prediction. Dating of the paintings was the most difficult task for all participants, particularly for novices (Fig. 3(B)). Familiarity levels were low and comparable for all images. Our stimulus selection can be therefore regarded as adequate (Fig. 2). Participants were more confident in their answers about the paintings they were more familiar with. In general, experts were more confident (Table 1). Confidence was also higher for correct answers and highest for the medium detection task. In order to explore task-dependent viewing strategies of our participants, we examined the eye movement parameters including fixation durations and saccade amplitudes. As expected, we found task-induced differences in fixation durations and saccade amplitudes. Longest fixation durations were found in the medium detection trials (Fig. 4(A)). In other words, detecting what

paint was used for the painting, whether it was oil, watercolor, pastel, etc. resulted in longer fixations compared to the two other tasks. In addition to longer fixations, the medium task induced shorter saccade amplitudes, however non-experts decreased their saccade amplitudes to a slightly larger extent, compared to experts during the same task (Fig. 4(B)). We also examined the similarity of fixation patterns of experts and non-experts in the different tasks. Our results showed overall less coherence and higher dispersion in fixation locations of experts compared to non-experts, indicating greater image coverage by experts. We also found less coherence and higher dispersion

in fixation locations for the task medium (Fig. 5). This indicates that while participants tried to detect medium of the paintings, the distributions of fixations they made were more diverse compared to two other tasks. This difference regarding medium task can be explained by the assumption, that informative areas for the medium task were more evenly distributed across the painting and allowed spatially more diverse fixations. Altogether, our findings imply that observers adjusted how they searched for information in the paintings depending on the task.

Generally, experts showed larger amplitudes across all three tasks (Table 1) and fixated larger areas of the paint-

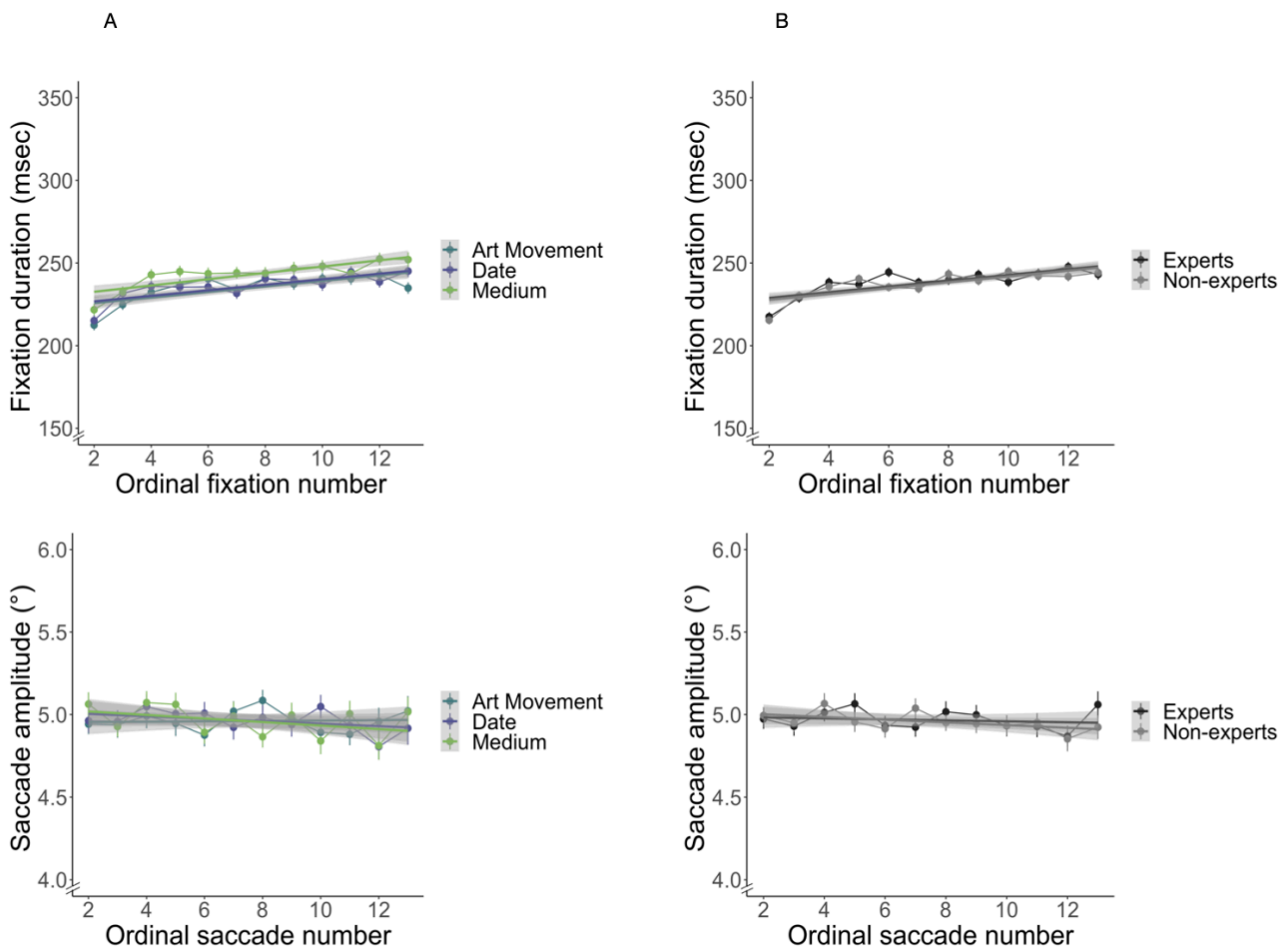


Figure 6: (A) Intergroup fixation density correlations for three tasks, (B) correlations between tasks for experts and non-experts.

ings. These findings are in line with previous studies supporting more holistic scanning in art experts (e.g. Fedorovskaya et al., 2017; Zangemeister et al., 1995) and are compatible with the results reported by Pihko et al. (2011) who observed larger distance of fixations from the image center in experts.

Temporal analysis of the trial showed the pattern of fixations becoming longer and saccades becoming smaller over the trial, regardless of task or expertise (Fig. 6). These results are in line with a well-established observation that fixation durations gradually increase after stimulus onset, while saccade amplitudes decrease (Antes, 1974; Buswell, 1935; Unema et al., 2005).

How our overall findings can be interpreted within a broader context of expertise guided visual search? If we think about theories from Gegenfurtner et al.'s (2011) meta-analysis, our result of larger saccade amplitudes in the expert group supports the first assumption of Kundel et al.'s (2007) holistic model of image perception. The second assumption of the model is that experts are also first to fixate task-relevant regions. This assumption could not be tested in this study.

According to the model, experts are able to rapidly extract a global scene impression during the initial scan and compare this impression with the prior knowledge about the image category. After this processing, experts need less time to first fixate task-relevant area. The model assumes that during the initial global scan, experts, due to their visual experience, are better at extracting information from their parafoveal and peripheral vision. The initial ambient mode of processing is then replaced with the focal mode of processing, where information is examined in more detail (for review see Sheridan & Reingold, 2017).

We know that visual expertise is domain specific and is based on a very specific knowledge of concrete image category. There is a huge amount of mixed results on expert fixation durations and saccade amplitude sizes depending on the visual domain (Brams et al., 2019; Gegenfurtner et al., 2011). However, generalizations about expertise within a domain should also be made with caution, as we have seen from our results, eye movement parameters very much depend on stimuli and difficulty of the task. This variability in stimuli and tasks might be the reason for the mixed results regarding fixations and amplitudes across the art domain. This work was the first attempt to examine task influences on art viewing. Future research

should shed more light on what stimuli and task characteristics constitute to the differences across visual domains.

Limitations

Our work clearly has some limitations. The stimuli set used in this study consisted of only 36 paintings covering only six art periods. A relatively low number of paintings, compared to other studies, for example, Wallraven et al. (2009) who selected 550 paintings covering 11 art periods for their style categorization study. Additionally, our artist coverage was biased. The stimulus set consisted of 36 paintings by 30 painters, with four paintings belonging to one painter. Although, the low familiarity rates and superior expert accuracy validates our choice of paintings, we believe that the balancing of stimulus set features, always has room for improvement. For example, due to the fact that we only used six art periods, the distribution of the art periods across art history timeline was not perfect. Skipping art periods between Baroque and Romanticism and periods between Romanticism and Impressionism led to gaps between years in the answer options. Conceivably, this could facilitate discrimination of some art movements compared to others according to the answer options.

Another controversial issue is whether students of art university departments fulfil the criteria of art expertise, the problem also discussed in Bauer and Schwan (2018). According to Ericsson and Lehmann (1996), experts in most domains attain their highest performance levels after a decade of intensive training. Although, the level of art expertise in our expert group was clearly higher than in the control group, nearly half of our experts were bachelor students in their second or third year of study, corresponding at most to an intermediate level of expertise. We assume that if we had an opportunity to test well-experienced art experts with prolonged training history, we could find even more profound differences between two groups.

Another limitation of our study, often discussed in the laboratory studies of art (e.g. Pihko et al., 2011), is that digitized paintings obviously do not have the same qualities as the original paintings. Whether it is sufficient to use digitized copies of paintings usually depends on the question of the study. For instance, it can be problematic to use digitized copies when measuring affective responses to paintings (Hayn-Leichsenring, 2017; see also Brieber et al., 2014). In our case, since conservator-restorers and art experts usually work with the original paintings while

solving the same tasks we gave our participants, we probably deprived our subjects of having some additional textural cues. Different physical-reflection properties on the surface of original paintings might be helpful for low-level feature extraction from the paintings.

It has been shown that liking and aesthetic preference is associated with longer fixation durations (e.g. Goller et al., 2019), probably also for non-biological object categories relevant for our sample of stimuli. However, since paintings were equally often assigned to three tasks and two groups, aesthetic differences between paintings can be disregarded. In order to explore the interaction of aesthetic preferences and task demands it would be interesting to include aesthetic preference ratings in future task-dependent art research.

As already discussed, task-induced differences highly depend on the variability between the distribution of informative locations in our images. The reason why we could not find any differences between the tasks date and art movement might be that informative areas of these tasks were overlapping. Although, observers were less accurate in the dating task, it can be assumed that two tasks are interconnected, considering the fact that if one can recognize the style, approximate date can be estimated and vice versa. It is conceivable that by setting these two tasks we simply asked the same question differently. The difficulty of predefining informative areas of the paintings for our three tasks without further inspection can be regarded as a general limitation of art stimuli. There are no general rules in visual art that can guide the gaze of observer through consecutive relevant areas, similar to mnemonic guides in medicine (e.g. Airways, Bones, Cardiac silhouette, Diaphragm) that help routinizing the eye movements of physicians.

Outlook

Our findings are consistent with the general assumption that observers adjust their viewing behavior according to the task. Overall, our results lend support to the notion, assuming experts' superiority in extracting task-relevant information based on their knowledge of scene content and scene statistics. Further investigation is needed to determine whether experts in art are indeed better in extracting task-relevant information than novices. For this purpose, criteria for informativeness needs to be proposed. Relevant ROIs in each painting and for each task can be defined according to the criteria. This way, the second assumption of

the holistic model of image processing can also be tested. It would be interesting to determine whether art experts, similar to medical experts (e.g. Brams et al., 2019) are faster in fixating task-relevant information. Future research might apply the meaning maps approach (Henderson & Hayes, 2017) to examine how semantic meaning distribution varies as a function of a task and what is the role of semantic representations in attention guidance in art viewing.

Even though viewing of art undoubtedly underlies human plans and task demands, this study is only the first step towards enhancing our understanding of task influences during art viewing. In addition to improving our understanding of task and expertise effects on eye movements, studies on art can give us some insights about art perception, sometimes with practical implications. For example, expertise-novice comparison can provide valuable information for the question of how to teach visual literacy, i.e. the ability to find meaning in imagery and efficient viewing strategies at art schools. Eye movement analysis generally, provides limitless possibilities of examining art history theories, artistic idiosyncrasies, exhibition concepts and many more. Findings from eye-tracking studies can lead to new ideas for enhancing visitor experiences in museums. As the quality of mobile eye-tracking improves, it certainly will accelerate research in the museum context, increasing ecological validity of the studies (e.g. Pelowski et al., 2018; Santini et al., 2018; Walker et al., 2017).

Conclusion

To sum up, we were able to show that, as in other known domains, task clearly influences eye movements during art viewing. We found task-induced differences in fixation durations and saccade amplitudes. Longest fixations and shortest saccade amplitudes were observed when art viewers were looking for cues to detect what paint material was used in the painting. These results can be interpreted in favor of observers' implemented task-specific strategy. It can be assumed that effortlessly detectable informative areas, in case of medium detection task, induced greater fixation variability and higher dispersion among observers, facilitated focused processing and ensured satisfactory task accuracy. With regard to expertise, we found less coherent fixation patterns among experts, suggesting more diversity in task-solving in this group. Our data also illustrates the effect of expertise on average saccade amplitude, supporting expertise-related holistic processing theory.

Ethics and Conflict of Interest

The author(s) declare(s) that the contents of the article are in agreement with the ethics described in <http://biblio.unibe.ch/portale/elibrary/BOP/jemr/ethics.html> and that there is no conflict of interest regarding the publication of this paper.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project number 222641018 – SFB/TRR 135 TP B2. Data and stimuli are available at: <http://doi.org/10.5281/zenodo.3956829>.

References

- Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, *103*(1), 62–70. <https://doi.org/10.1037/h0036799>
- Augustin, M. D., & Leder, H. (2006). Art expertise: A study of concepts and conceptual spaces. *Psychology Science*, *48*(2), 135–156.
- 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>
- Bakdash, J. Z., & Marusich, L. R. (2017). Repeated Measures Correlation. *Frontiers in Psychology*, *8*. <https://doi.org/10.3389/fpsyg.2017.00456>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). <https://doi.org/10.18637/jss.v067.i01>
- Bauer, D., & Schwan, S. (2018). Expertise influences meaning-making with Renaissance portraits: Evidence from gaze and thinking-aloud. *Psychology of Aesthetics, Creativity, and the Arts*, *12*(2), 193–204. <https://doi.org/10.1037/aca0000085>
- Borji, A., & Itti, L. (2014). Defending Yarbus: Eye movements reveal observers' task. *Journal of Vision*, *14*(3), 29–29. <https://doi.org/10.1167/14.3.29>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. <https://doi.org/10.1163/156856897X00357>
- Brams, S., Ziv, G., Levin, O., Spitz, J., Wagemans, J., Williams, A. M., & Helsen, W. F. (2019). The relationship between gaze behavior, expertise, and performance: A systematic review. *Psychological Bulletin*, *145*(10), 980–1027. <https://doi.org/10.1037/bul0000207>
- Brieber, D., Nadal, M., Leder, H., & Rosenberg, R. (2014). Art in Time and Space: Context Modulates the Relation between Art Experience and Viewing Time. *PLoS ONE*, *9*(6), e99019. <https://doi.org/10.1371/journal.pone.0099019>
- Buswell, G. T. (1935). *How people look at pictures: a study of the psychology and perception in art*. University of Chicago Press.
- Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*(3), 6–6. <https://doi.org/10.1167/9.3.6>
- Cerf, M., Frady, E. P., & Koch, C. (2009). Faces and text attract gaze independent of the task: Experimental data and computer model. *Journal of Vision*, *9*(12), 10–10. <https://doi.org/10.1167/9.12.10>
- Chun, M. M., & Jiang, Y. (1998). Contextual Cueing: Implicit Learning and Memory of Visual Context Guides Spatial Attention. *Cognitive Psychology*, *36*(1), 28–71. <https://doi.org/10.1006/cogp.1998.0681>
- Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The EyeLink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 613–617. <https://doi.org/10.3758/bf03195489>
- Cronin, D. A., Hall, E. H., Goold, J. E., Hayes, T. R., & Henderson, J. M. (2020). Eye Movements in Real-World Scene Photographs: General Characteristics and Effects of Viewing Task. *Frontiers in Psychology*, *10*. <https://doi.org/10.3389/fpsyg.2019.02915>
- DeAngelus, M., & Pelz, J. B. (2009). Top-down control of eye movements: Yarbus revisited. *Visual Cognition*, *17*(6–7), 790–811. <https://doi.org/10.1080/13506280902793843>

- Dorr, M., Martinetz, T., Gegenfurtner, K. R., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision, 10*(10), 28–28. <https://doi.org/10.1167/10.10.28>
- Einhäuser, W., Rutishauser, U., & Koch, C. (2008a). Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *Journal of Vision, 8*(2), 2. <https://doi.org/10.1167/8.2.2>
- Einhäuser, W., Spain, M., & Perona, P. (2008b). Objects predict fixations better than early saliency. *Journal of Vision, 8*(14), 18–18. <https://doi.org/10.1167/8.14.18>
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211–245. <https://doi.org/10.1037/0033-295X.102.2.211>
- Ericsson, K. A., & Lehmann, A. C. (1996). EXPERT AND EXCEPTIONAL PERFORMANCE: Evidence of Maximal Adaptation to Task Constraints. *Annual Review of Psychology, 47*(1), 273–305. <https://doi.org/10.1146/annurev.psych.47.1.273>
- Fedorovskaya, E., Kapisthalam, S., & Bu, Y. (2017). Gaze Patterns in Art Viewing and their Dependencies on Expertise and Image Characteristics. *Electronic Imaging, 2017*(14), 257–260. <https://doi.org/10.2352/ISSN.2470-1173.2017.14.HVEI-153>
- Fudali-Czyż, A., Francuz, P., & Augustynowicz, P. (2018). The Effect of Art Expertise on Eye Fixation-Related Potentials During Aesthetic Judgment Task in Focal and Ambient Modes. *Frontiers in Psychology, 9*. <https://doi.org/10.3389/fpsyg.2018.01972>
- Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise Differences in the Comprehension of Visualizations: a Meta-Analysis of Eye-Tracking Research in Professional Domains. *Educational Psychology Review, 23*(4), 523–552. <https://doi.org/10.1007/s10648-011-9174-7>
- Goller, J., Mitrovic, A., & Leder, H. (2019). Effects of liking on visual attention in faces and paintings. *Acta Psychologica, 197*, 115–123. <https://doi.org/10.1016/j.actpsy.2019.05.008>
- Gombrich, E. H. (1966). *The Story of Art*. Phaidon Publishers.
- Haider, H., & Frensch, P. A. (1999). Eye movement during skill acquisition: More evidence for the information-reduction hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(1), 172–190. <https://doi.org/10.1037/0278-7393.25.1.172>
- Hayhoe, M. (2000). Vision Using Routines: A Functional Account of Vision. *Visual Cognition, 7*(1–3), 43–64. <https://doi.org/10.1080/135062800394676>
- Hayn-Leichsenring, G. U. (2017). The Ambiguity of Artworks – A Guideline for Empirical Aesthetics Research with Artworks as Stimuli. *Frontiers in Psychology, 8*. <https://doi.org/10.3389/fpsyg.2017.01857>
- Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*(11), 498–504. <https://doi.org/10.1016/j.tics.2003.09.006>
- Henderson, J. M., & Hayes, T. R. (2017). Meaning-based guidance of attention in scenes as revealed by meaning maps. *Nature Human Behaviour, 1*(10), 743–747. <https://doi.org/10.1038/s41562-017-0208-0>
- Holmqvist, K., & Andersson, R. (2017). *Eye tracking: A comprehensive guide to methods, paradigms and measures*. Lund Eye-Tracking Research Institute.
- Ishihara, S. (1960). *Tests for colour-blindness*. Kanehara Shuppan Company.
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194–203. <https://doi.org/10.1038/35058500>
- Koide, N., Kubo, T., Nishida, S., Shibata, T., & Ikeda, K. (2015). Art Expertise Reduces Influence of Visual Saliency on Fixation in Viewing Abstract-Paintings. *PLOS ONE, 10*(2), e0117696. <https://doi.org/10.1371/journal.pone.0117696>
- Kristjanson, A. F., & Antes, J. R. (1989). Eye movement analysis of artists and nonartists viewing paintings. *Visual Arts Research, 15*(2), 21–30. <https://www.jstor.org/stable/20715702?seq=1>
- Kundel, H. L., Nodine, C. F., Conant, E. F., & Weinstein, S. P. (2007). Holistic Component of Image Perception in Mammogram Interpretation: Gaze-tracking Study. *Radiology, 242*(2), 396–402. <https://doi.org/10.1148/radiol.2422051997>

- Land, M., Mennie, N., & Rusted, J. (1999). The Roles of Vision and Eye Movements in the Control of Activities of Daily Living. *Perception*, 28(11), 1311–1328. <https://doi.org/10.1068/p2935>
- Land, M. F., & Lee, D. N. (1994). Where we look when we steer. *Nature*, 369(6483), 742–744. <https://doi.org/10.1038/369742a0>
- Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, 95(4), 489–508. <https://doi.org/10.1348/0007126042369811>
- Loon, E. M. V., Hooge, I. T. C., & Berg, A. V. V. d. (2002). The timing of sequences of saccades in visual search. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269(1500), 1571–1579. <https://doi.org/10.1098/rspb.2002.2062>
- Massaro, D., Savazzi, F., Di Dio, C., Freedberg, D., Gallese, V., Gilli, G., & Marchetti, A. (2012). When Art Moves the Eyes: A Behavioral and Eye-Tracking Study. *PLoS ONE*, 7(5), e37285. <https://doi.org/10.1371/journal.pone.0037285>
- Mills, M., Hollingworth, A., Van der Stigchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision*, 11(8), 17–17. <https://doi.org/10.1167/11.8.17>
- Molnar, F. (1981). About the Role of Visual Exploration in Aesthetics. In: Day H.I. (Eds.) *Advances in Intrinsic Motivation and Aesthetics*. Springer.
- Nodine, C. F., Locher, P. J., & Krupinski, E. A. (1993). The role of formal art training on perception and aesthetic judgement of art compositions. *Leonardo*, 26(3), 219–227. <https://doi.org/10.2307/1575815>
- Nuthmann, A. (2016). Fixation durations in scene viewing: Modeling the effects of local image features, oculomotor parameters, and task. *Psychonomic Bulletin & Review*, 24(2), 370–392. <https://doi.org/10.3758/s13423-016-1124-4>
- Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational model of fixation durations in scene viewing. *Psychological Review*, 117(2), 382–405. <https://doi.org/10.1037/a0018924>
- Over, E. A. B., Hooge, I. T. C., Vlaskamp, B. N. S., & Erkelens, C. J. (2007). Coarse-to-fine eye movement strategy in visual search. *Vision Research*, 47(17), 2272–2280. <https://doi.org/10.1016/j.visres.2007.05.002>
- Pang, C. Y., Nadal, M., Müller-Paul, J. S., Rosenberg, R., & Klein, C. (2013). Electrophysiological correlates of looking at paintings and its association with art expertise. *Biological Psychology*, 93(1), 246–254. <https://doi.org/10.1016/j.biopsycho.2012.10.013>
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42(1), 107–123. [https://doi.org/10.1016/s0042-6989\(01\)00250-4](https://doi.org/10.1016/s0042-6989(01)00250-4)
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies, *Spatial Vision*, 10, 437–442. <https://doi.org/10.1163/156856897X00366>
- Pelowski, M., Leder, H., Mitschke, V., Specker, E., Genger, G., Tinio, P. P. L., Vaporova, E., Bieg, T., & Husslein-Arco, A. (2018). Capturing Aesthetic Experiences with Installation Art: An Empirical Assessment of Emotion, Evaluations, and Mobile Eye Tracking in Olafur Eliasson’s “Baroque, Baroque!” *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.01255>
- Peters, R. J., Iyer, A., Itti, L., & Koch, C. (2005). Components of bottom-up gaze allocation in natural images. *Vision Research*, 45(18), 2397–2416. <https://doi.org/10.1016/j.visres.2005.03.019>
- Pihko, E., Virtanen, A., Saarinen, V., Pannasch, S., Hirvenkari, L., Tossavainen, T., Haapala, A., & Hari, R. (2011). Experiencing Art: The Influence of Expertise and Painting Abstraction Level. *Frontiers in Human Neuroscience*, 5. <https://doi.org/10.3389/fnhum.2011.00094>
- Rayner, K. (2009). 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>
- Santini, T., Brinkmann, H., Reitstätter, L., Leder, H., Rosenberg, R., Rosenstiel, W., & Kasneci, E. (2018). The art of pervasive eye tracking. Proceedings of the 7th Workshop on Pervasive Eye Tracking and Mobile Eye-Based Interaction - PETMEI '18. the 7th Workshop. <https://doi.org/10.1145/3208031.3208032>

- Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision*, 11(5), 9–9. <https://doi.org/10.1167/11.5.9>
- Sheridan, H., & Reingold, E. M. (2017). The Holistic Processing Account of Visual Expertise in Medical Image Perception: A Review. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01620>
- Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, 7(14), 4. <https://doi.org/10.1167/7.14.4>
- Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision*, 11(5), 5–5. <https://doi.org/10.1167/11.5.5>
- Tatler, B. W., Wade, N. J., Kwan, H., Findlay, J. M., & Velichkovsky, B. M. (2010). Yarbus, Eye Movements, and Vision. *I-Perception*, 1(1), 7–27. <https://doi.org/10.1068/i0382>
- Theeuwes, J. (2018). Visual Selection: Usually Fast and Automatic; Seldom Slow and Volitional. *Journal of Cognition*, 1(1). <https://doi.org/10.5334/joc.13>
- Theeuwes, J. (2019). Goal-driven, stimulus-driven, and history-driven selection. *Current Opinion in Psychology*, 29, 97–101. <https://doi.org/10.1016/j.copsyc.2018.12.024>
- Unema, P. J. A., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, 12(3), 473–494. <https://doi.org/10.1080/13506280444000409>
- Võ, M. L. H., & Henderson, J. M. (2009). Does gravity matter? Effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *Journal of Vision*, 9(3), 24–24. <https://doi.org/10.1167/9.3.24>
- Vogt, S., & Magnussen, S. (2007). Expertise in Pictorial Perception: Eye-Movement Patterns and Visual Memory in Artists and Laymen. *Perception*, 36(1), 91–100. <https://doi.org/10.1068/p5262>
- Walker, F., Bucker, B., Anderson, N. C., Schreij, D., & Theeuwes, J. (2017). Looking at paintings in the Vincent Van Gogh Museum: Eye movement patterns of children and adults. *PLOS ONE*, 12(6), e0178912. <https://doi.org/10.1371/journal.pone.0178912>
- Wallraven, C., Fleming, R., Cunningham, D., Rigau, J., Feixas, M., & Sbert, M. (2009). Categorizing art: Comparing humans and computers. *Computers & Graphics*, 33(4), 484–495. <https://doi.org/10.1016/j.cag.2009.04.003>
- Wallraven, C., Kaulard, K., Kürner, C., & Pepperell, R. (2008). In the Eye of the Beholder: The Perception of Indeterminate Art. *Leonardo* 41(2), 116–117. <https://www.muse.jhu.edu/article/236373>
- Williams, C. C., & Castelano, M. S. (2019). The Changing Landscape: High-Level Influences on Eye Movement Guidance in Scenes. *Vision*, 3(3), 33. <https://doi.org/10.3390/vision3030033>
- Yarbus, A. L. (1967). *Eye Movements and Vision*. <https://doi.org/10.1007/978-1-4899-5379-7>
- Ylitalo, A.-K., Särkkä, A., & Guttorp, P. (2016). What we look at in paintings: A comparison between experienced and inexperienced art viewers. *The Annals of Applied Statistics*, 10(2), 549–574. <https://doi.org/10.1214/16-aos921>
- Zangemeister, W. H., Sherman, K., & Stark, L. (1995). Evidence for a global scanpath strategy in viewing abstract compared with realistic images. *Neuropsychologia* 33, 1009–1025. [https://doi.org/10.1016/0028-3932\(95\)00014-T](https://doi.org/10.1016/0028-3932(95)00014-T)