

How words can guide our eyes

Increasing engagement with art through audio-guided visual search in young and older adults

Naomi Vingron,¹ Lea Alexandra Müller Karoza,¹
Nancy Azevedo,² Aaron Johnson,³
Evdokimos Konstantinidis,⁴ Panagiotis Bamidis,⁴ Melissa Vö,¹
and Eva Kehayia^{2,5}

¹Goethe University | ²Center for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR – Jewish Rehabilitation Hospital (JRH) of the CISSS de Laval) | ³Concordia University |

⁴Aristotle University | ⁵McGill University

Pursuing cognitively stimulating activities, such as engaging with art, is crucial to a healthy lifestyle. The current work simulates visits to an art museum in a laboratory setting. Using eye tracking, we explored how linguistically guided visual search may increase attention, enjoyment and retention of information when viewing art. Two groups of adults, young (under 35 years) and older (over 65 years) viewed ten paintings on a computer screen presented either with or without an accompanying audio-guide, while having their eye movements recorded. Audio-guides referred to specific areas of the painting, marked as Interest Areas (IA). Across age groups, as attested by gaze fixations, the audio-guides increased attention to these areas compared to free-viewing. Audio-guided viewing did not lead to a significantly increase over free-viewing in information recall accuracy or feelings of enjoyment and engagement. Overall, older adults did report feeling more positively about both audio-guided and free viewing than young adults. Thus, the use of audio-guides, specifically the gamification through linguistically guided visual search, may be a useful tool to promote meaningful attentional interactions with art.

Keywords: visual search, eye movements, museum, audio-guide

Introduction

Engaging in cognitively stimulating activities such as a visit to the museum is crucial to the maintenance of whole person health throughout the lifespan. While several scientific studies point to the positive effects of various psychosocial activities on subjective wellbeing (e.g., Cohen-Mansfield, 2005; Solway et al., 2015), such activities can be challenging for many, particularly individuals with language and other cognitive impairments. In order to develop tools that promote such activities, both the desired outcome of the visit overall (i.e., enjoyment, return visits, etc.; Anderson et al., 2007; Hutchinson & Eardley, 2020) and the psycholinguistic processes unfolding a museum visit must be considered. The present study simulates visits to an art museum in a laboratory setting using audio-guides presenting easy-to-understand information and encouraging viewers to engage with the painting through guided visual search.

Knowing where to direct one's attention in a museum environment is challenging, despite the availability of information intended to facilitate understanding of the message of the exhibit. The challenge is twofold: exhibits typically present a large array of visual stimuli that may overwhelm the visitor (Bitgood, McKerchar & Dukes, 2013) and linguistic information (i.e., text on the wall, audio-guides) often uses complex language making it unhelpful or even inaccessible to older individuals experiencing language limitations and language processing difficulties (Brien, submitted). In such situations, museum visitors tend to engage in a *browsing behavior*, where they spend only a very short time in front of any one artwork (Smith & Smith, 2001). To address this, *the attention-value model* (Bitgood, McKerchar & Dukes, 2013) suggests that visitors pay attention to an exhibition if they perceive high value in the encounter. Importantly, value is more than just enjoyment – it also includes an analytic and educational aspect, which requires background information presented in an accessible manner (e.g., one may not find a painting beautiful, yet find the context it was created in may be interesting enough to feel engaged). Thus, accessible linguistic information guiding the viewing process may act as the bridge between enjoyment and engagement with art.

Engaging older visitors' attention to produce an implicit learning outcome requires a fine balance between providing meaningful information without being too effortful. As work involving the visual world and visual search experimental paradigms has shown, attentional guidance to a target occurs largely automatically and memory for said target is improved as interactions with it become more meaningful (i.e., context) and engaging (Draschkow et al., 2014; Helbing et al., 2022; Wolfe, 1994). Our experimental paradigm aims to achieve this balance of meaning and effort through audio-guides that use accessible language guiding

the visitors to target their attention in a way that will lead to understanding the description.

Thinking about what museum visitors are already accustomed to, audio-guides constitute an ideal medium, allowing for carefully chosen spoken words to facilitate a more accessible form of engagement. Not only is the auditory presentation of information more accessible to many, but museum audio-guides have been linked to more meaningful attentional interactions and improved information processing among individuals of diverse ages and cognitive abilities (Bauer-Krösbacher, 2013; Hutchinson & Eardley, 2020). However, few guidelines exist on the creation of linguistic and cognitive accessibility support tools.

Current study

The current work simulates visits to an art museum in a laboratory setting, using gamification (Deterding, Dixon, Khaled & Nacke, 2011) to encourage more in-depth interactions with the art pieces. Eye-movements will be recorded as participants view pieces of art freely or while listening to an audio-guide. More specifically, we developed audio-guides that present easy-to-understand information about the art as well as encourage the viewer to engage with the piece through accessible descriptions of the art (e.g., statements such as: “In this painting, the children are sitting around the table. The items in their hands and on the table suggest that it is dinner time”).

This easily expandable and reproducible framework primarily targets visual attention, but also encompasses other mental functions including listening, and creativity. Our aim is to establish that (a) participants’ eye movements reflect engagement with the task overall and attention to specific parts of the paintings (i.e., performing the linguistically guided visual search task), that (b) there is an immediate link between an unfolding utterance about regions of a painting and participants’ attention being guided there and finally that (c) engaging in this activity will lead greater engagement and enjoyment of the overall art viewing experience. To this end, we asked a control group of young adults and a group of healthy older adults to view paintings that were either accompanied by an audio-guide or not, while their eye movements were recorded.

Methods

Stimuli

We selected 10 paintings from the Montreal Fine Arts Museum, the Quebec Fine Arts Museum and the Städel Museum in Frankfurt am Main. All depicted scenes that included humans, animals and/ or objects.

Following recommendations of the Brien et al. (2024 forthcoming) study at the Montreal Museum of Fine Arts, where linguistically accessible information on paintings was much preferred by older adults, audio-guides were prepared at about a 9th grade reading level (using simple and frequent words in short sentences) in German.¹ They were recorded by native speakers of the language and played through computer speakers. Audio-guides were between 74–101 seconds long. Each audio-guide referred to three Interest Areas (IAs; see Figure 1). The complete set of paintings and accompanying audio-guides may be found on the OSF repository (https://osf.io/tzy3p/?view_only=f983331b65a74ece8ef a48499ed361cd).



Figure 1. Example of painting with critical utterances referring to IAs

Procedure

Participants were randomly assigned to either a free-viewing or an audio-guided viewing condition, in which accompanying audio-guides referred to areas of the painting they were viewing. Paintings were displayed for the exact duration of the audio-guide in both the guided and free-viewing conditions. The paintings

1. Brien et al. (2024 forthcoming) had determined that the average reading grade level of exhibit texts is 14.6 for French texts and 13.33 for English texts, corresponding to 13–14 years of education completed in that language.

were presented on a 24-in. monitor with a refresh rate of 128 Hz. Participants were positioned in a chinrest 65 cm from the screen, and eye movements were recorded with the Eyelink-1000+ desktop mount (SR Research, Ontario, Canada) at 500 Hz.

After viewing each of the 10 paintings, participants were asked content and visual recall questions. The number of questions differed between the young and older adult groups, with the older adults always being asked two additional questions to account for this group possibly experiencing more difficulties using the keyboard to make their responses.² The goal of asking these questions was to ensure a baseline level of engagement.

Subsequently, all participants completed a short questionnaire about their feelings of enjoyment and engagement during the task. The questionnaire, administered through Google Forms, presented statements along with a 10-point Likert scale (1 – strongly disagree to 10 – strongly agree). A copy of the questionnaires may be found on the OSF repository (https://osf.io/tzy3p/?view_only=f983331b65a74ece8efa48499ed361cd).

Participants

To establish baseline viewing and attentional patterns, we first asked thirteen young adults (ten females and three males; mean age = 20.91; $SD = 2.10$) to view ten paintings on a computer screen without an audio-guide or any explicit instructions (i.e., free-viewing task). A group of 16 older adults (seven females and nine males; mean age = 71.62; $SD = 5.01$) also performed the same task. We then investigated the use of audio-guides, in two separate groups: one of 13 young adults (eight females and five males; mean age = 20.92; $SD = 2.43$) and a group of 16 older adults (nine females and seven males; mean age = 72.00; $SD = 5.53$). All participants reported having normal or corrected to normal vision and hearing.

As the task was administered entirely in German, we made sure that German was the most dominant language for all participants. At the time of the study, all young adults were pursuing undergraduate degrees and most of the older adult participants were enrolled in lifelong learning courses at the university (*Universität des 3. Lebensalters*) while others had at least some education beyond high school.

2. Specifically, in the free-viewing condition, young adult participants were asked one content recall question about what was visible in the painting (true/false), while older adults were asked three such questions. In the audio-guided condition, both groups were asked two additional content recall questions pertaining to information mentioned in the audio-guide (YA: three questions in total; OA: five questions in total). Additionally, across conditions both groups were asked one visual recall question in which two puzzle pieces were presented and participants had to decide which one had been part of the painting (two-alternative forced-choice).

Young adult participants were recruited from an experimental practicum course at Goethe University in Frankfurt and through word of mouth. Older adult participants were recruited through Goethe University's lifelong learning program as well as flyers and word of mouth. Participants received 10 €/ hour for their participation or where applicable, course credit. Of note, young adults took around 30 minutes to complete the study, while older adults needed an hour, as both calibration and response times were longer for this age group.

Results

Analysis

Areas of the painting that were referenced by the audio-guided visual search were marked as interest areas (IAs). As well, the time point when these areas were mentioned in the audio description was cross-referenced with the eye movements recorded to assess whether participants look at the areas being described.

We conducted three separate analyses. Firstly, we compared self-reported feelings of enjoyment and engagement of the free-viewing and audio-guided tasks across age groups. Secondly, we compared general patterns of attention to IAs (i.e., visual search targets) when attention was explicitly guided to them (audio-guide condition) vs. when the painting was viewed freely (i.e., inherent interest). In this analysis a Data Viewer (SR Research Ltd., version 4.3.1) interest area report was used. Finally, we examined the attentional patterns on IAs over time; that is, fixations to those areas while they are being mentioned in the audio-guide. We achieved this using the Visual World Pre package in R (version 1.2.4; Porretta et al., 2018) to preprocess the sample report data retrieved from Data Viewer. Down-sampling to 40 Hz allowed us to conduct a time-course analysis and generate figures.

Of note, overall recall accuracy was high ($M=83.75\%$) and did not differ across experiments (see Appendix 1).

Analysis 1: Self-reported engagement and enjoyment

We first compared self-reported engagement and enjoyment of the tasks across groups (see Appendix 1). A generalized linear model revealed no difference in self-reported engagement with the task across experimental conditions, however, OA reported significantly higher overall engagement ($p < .05$). A separate model revealed a marginal increase in task enjoyment with an audio-guide compared to the free-viewing condition ($p = .09$). As well, OA reported more enjoyment of the tasks overall ($p < .05$).

Analysis 2: Overall proportion of looks to interest areas (IAs)

To the extent that visual attention can be guided by auditory input, listening to an audio-guide while viewing a painting should increase the number and duration of fixations to areas that are explicitly mentioned. That is, additional, linguistically presented information may make specific areas of a painting more interesting than if a viewer was freely scanning. To test this, we compared eye movement patterns, specifically the proportion of fixations (i.e., how frequently a person looks at a specific IA) and dwell times (i.e., the total amount of time spent looking at a specific IA) across the audio-guided and free-viewing conditions for older and young adults (see Table 1 and Figure 2). Importantly, this analysis considers eye movements throughout the entire viewing period.

A generalized linear model for proportion of fixations on the IA with the variables group (OA and YA) and experiment (free-viewing and audio-guided) serving as predictors revealed no significant effects. A subsequent linear model for IA dwell time revealed that both YA and OA participants fixated IAs mentioned in the audio-guide condition significantly longer compared to the free-viewing condition ($p < .05$). Dwell times were also marginally longer among young adults compared to older adults ($p = 0.07$).

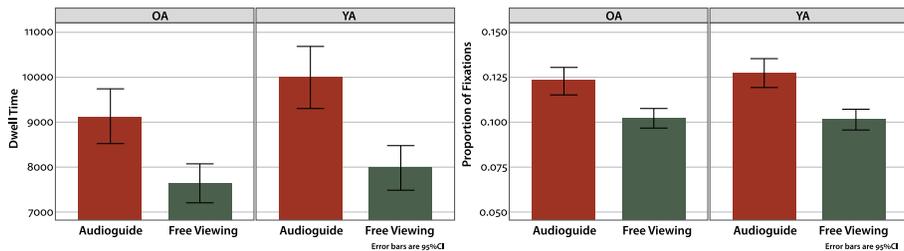


Figure 2. Left: Time allocated to IAs across groups and experiments. Right: Proportion of looks to IAs across groups and experiments

Analysis 3: Time course analysis of proportion of looks to interest areas (IAs)

We hypothesize that while an area is being referred to in the audio-guide, the number of fixations on this area will increase in contrast to all other IAs. To test this hypothesis, we collapsed observations across paintings and compared the proportion of fixations on one of the three IAs as it is being described in the audio-guide to the remaining two IAs (i.e., the IAs not being currently being mentioned are effectively random areas). Importantly, we will only look at this in the audio-guide experiment. We consider only fixations made while the audio-guide refers to an area. As the resulting time window varied for each painting and IA, we settled on using a 550-millisecond time window to illustrate the results in

Figure 3 below but considered the full duration of each IA being mentioned for the generalized linear model (also see tables in Appendix 2).

The generalized linear models run for each IA separately all revealed that while the respective IA was being mentioned, it received significantly more fixations than either of the other regions identified as IAs at different points in the audio-guide (all $p < .05$). Further, YA made more fixations to the IA than OA ($p < .05$).

Discussion

The results of our study are promising with regards to the use of linguistically guided visual search to modulate attentional patterns during art viewing. Our main hypothesis was that eye movements reflect participants' engagement with the task in general and their attention to specific parts of the paintings. In fact, we found a close link between linguistic cues and attention allocation both across the entire viewing period when comparing the audio-guided and free-viewing conditions as well as when examining the time course of visual attention during a critical utterance in the audio-guide condition. While audio-guides increased implicit engagement (i.e., eye movements), we didn't find that their use significantly increased self-reported engagement and enjoyment of the task. Of note, OA participants did report more engagement and enjoyment than YA did, suggesting a general benefit of task structure among OA.

Engaging with art can become more challenging as cognitive abilities decline with advancing age. Thus, we expected older adults to particularly benefit from accessible linguistic input to help them direct and focus their attention. We were able to show that audio-guides directed both older and young adult participants' attention in an immediate way. This suggests that using easy-to-understand words to guide attention led to increased implicit engagement with the art. However, the use of audio guides did not affect explicit feelings of enjoyment possibly due to the goal-directed nature of the task.

In sum, our findings suggest that linguistically guided visual search, delivered through accessible audio-guides, can be seuseful and constitute an important motivator for individuals of all ages to integrate museum activities in a sustained healthy and active lifestyle.

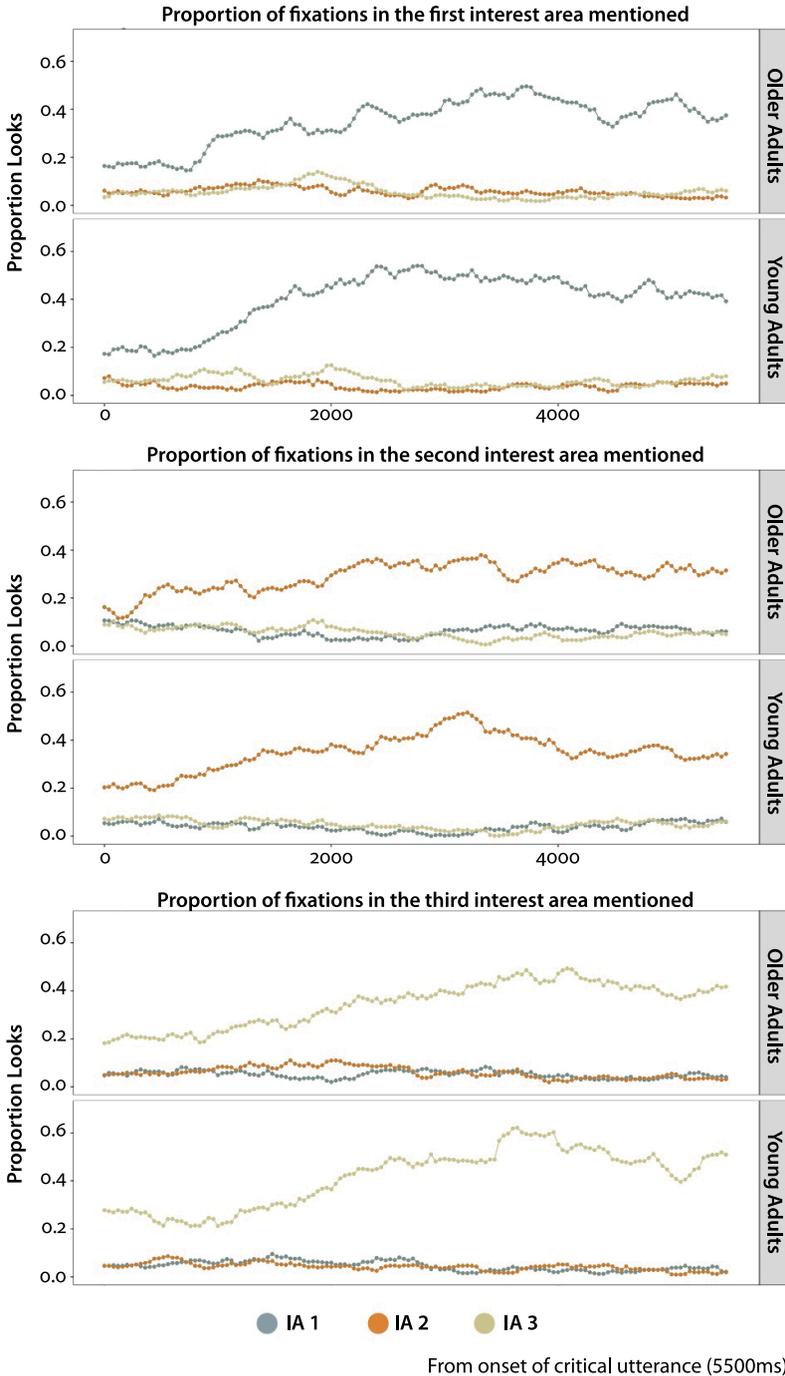


Figure 3. Proportions of fixations on IAs (i.e., visual search target) unfolding over 5500 ms from onset of critical utterances

Funding

Open Access publication of this article was funded through a Transformative Agreement with Goethe University Frankfurt.

References

- Anderson, D., Storcksdieck, M., & Spock, M. (2007). Understanding the long-term impacts of museum experiences. *In principle, in practice: Museums as learning institutions*, 197–215.
- Bauer-Krösbacher, C. (2013). Mobile interpretation at cultural attractions: insights into users and non-users of audio-guides. *In Cultural tourism* (pp. 64–73). Wallingford UK: CABI.
- Bitgood, S., McKerchar, T.L., & Dukes, S. (2013). Looking back at Melton: Gallery density and visitor attention. *Visitor Studies*, 16(2), 217–225.
- Brien, C., Kehayia, E., Jarema, G., Alary Gauvreau, C. Poldma, Giroux, L. & Duhaime, J. Improving Accessibility of Information at the Montreal Museum of Fine Arts Musées pour tous, 2024 forthcoming (eds : Porcedda, A., Kehayia, E., Hamouni, Z.)
- Cohen-Mansfield, J. (2005). Nonpharmacological interventions for persons with dementia. *Alzheimer's Care Today*, 6(2), 129–145.
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011, September). From game design elements to gamefulness: defining “gamification”. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (pp. 9–15).
- Draschkow, D., Wolfe, J.M., & Vo, M.L.H. (2014). Seek and you shall remember: Scene semantics interact with visual search to build better memories. *Journal of Vision*, 14(8), 10–10.
- Helbing, J., Draschkow, D., & L.-H. Vö, M. (2022). Auxiliary scene-context information provided by anchor objects guides attention and locomotion in natural search behavior. *Psychological Science*, 33(9), 1463–1476.
- Hutchinson, R. S., & Eardley, A. F. (2020). The accessible museum: towards an understanding of international audio description practices in museums. *Journal of Visual Impairment & Blindness*, 114(6), 475–487.
- Porretta, V., Kyröläinen, A. J., van Rij, J., & Järvikivi, J. (2018). Visual world paradigm data: From preprocessing to nonlinear time-course analysis. In *Intelligent Decision Technologies 2017: Proceedings of the 9th KES International Conference on Intelligent Decision Technologies (KES-IDT 2017)–Part II 9* (pp. 268–277). Springer International Publishing.
- Smith, J. K., & Smith, L. F. (2001). Spending time on art. *Empirical studies of the arts*, 19(2), 229–236.
- Solway, R., Thompson, L., Camic, P.M., & Chatterjee, H. J. (2015). Museum object handling groups in older adult mental health inpatient care. *International Journal of Mental Health Promotion*, 17(4), 201–214.
- Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic bulletin & review*, 1, 202–238.

Appendix 1. Average recall accuracy and questionnaire responses

	Content Recall Accuracy		Visual Recall Accuracy		Engagement (10-point scale)		Enjoyment (10-point scale)	
	YA	OA	YA	OA	YA	OA	YA	OA
Free-Viewing	72%	77%	95%	89%	8.11	8.58	6.3	8.37
Audio-Guide	81%	76%	93%	88%	7.61	8.91	7.25	8.72

Appendix 2. Mean proportions of fixations on the IA while it is being mentioned in the audio guide

	Proportion of Fixations					
	OA			YA		
	IA 1	IA 2	IA 3	IA 1	IA 2	IA 3
IA 1 being mentioned	34%	5%	5%	39%	3%	5%
IA 2 being mentioned	4%	29%	6%	3%	36%	5%
IA 3 being mentioned	4%	6%	35%	5%	4%	39%

Address for correspondence

Naomi Vingron
 Psychology Department
 Goethe University Frankfurt
 Theodor-W.-Adorno-Platz 6
 PEG #5.Go78
 60323 Frankfurt, am Main
 Germany
 vingron@psych.uni-frankfurt.de

Co-author information

Lea Alexandra Müller Karoza
Goethe University Frankfurt
muellerkaroza@psych.uni-frankfurt.de

Nancy Azevedo
Center for Interdisciplinary Research in
Rehabilitation of Greater Montreal (CRIR –
Jewish Rehabilitation Hospital (JRH) of the
CISSS de Laval)
nancy.azevedo.ciessler@sss.gouv.qc.ca

Aaron Johnson
Concordia University
aaron.johnson@concordia.ca

Evdokimos Konstantinidis
Aristotle University
evdokimosk@gmail.com

Panagiotis Bamidis
Aristotle University
bamidis@med.auth.gr

Melissa Vö
Goethe University
mlvo@psych.uni-frankfurt.de

Eva Kehayia
Center for Interdisciplinary Research in
Rehabilitation of Greater Montreal (CRIR –
Jewish Rehabilitation Hospital (JRH) of the
CISSS de Laval)
McGill University
eva.kehayia@mcgill.ca

Publication history

Date received: 15 April 2024

Date accepted: 28 November 2024

Published online: 10 January 2025