The Open Gallery for Arts Research (OGAR): An open-source tool for studying the psychology of virtual art museum visits

Rebekah M. Rodriguez-Boerwinkle[1] · Martin J. Boerwinkle[2] · Paul J. Silvia[1]

Abstract

To expand the tools available to arts researchers in psychology, we present the Open Gallery for Arts Research (OGAR), a free, open-source tool for studying visitor behavior within an online gallery environment. OGAR is highly extensible, allowing researchers to modify the environment to test different hypotheses, and it affords assessing a wide range of outcome variables. After describing the tool and its development, we present a proof-of-concept study that evaluates OGAR's usability and performance and illustrates some ways that it can be used to study the psychology of virtual visits. With a sample of 44 adults from an online participant panel who freely explored OGAR, we observed that OGAR had good usability based on high scores on the System Usability Scale and rare instances of self-reported nausea, among other usability markers. Furthermore, using position and viewing data provided by OGAR, we found that participants navigated the gallery and interacted with the artwork in predictable and coherent ways that resembled visitor behavior in real-world art museums. OGAR appears to be a promising tool for researchers and art professionals interested in how people navigate and experience virtual and real art spaces.

What were once private collections guarded by the societal elite, symbols of wealth and status, and a means of distinguishing between the "cultured" few and the "uncultured" many, art museums are now cultural institutions that aim to serve the masses (Bennett, 2013). With stated mission statements like the Metropolitan Museum of Art's— "to collect, preserve, study, exhibit, and stimulate appreciation for and advance knowledge of works of art"—museums now emphasize their roles as disseminators of knowledge and culture (Metropolitan Museum of Art mission statement, 2000). As part of this mission, the interdisciplinary study of the psychology of museum experiences—grounded in the psychology of the arts, visitor studies, and art education—seeks to understand how people experience, understand, and learn from their time spent in art museums (Tinio et al., 2015).

In the present research, we aim to expand the tools available to researchers in this growing scholarly field by developing the open gallery for arts research (OGAR). OGAR is a free, open-source tool for studying visitor behavior within an online gallery environment. It is highly extensible, allowing researchers to modify the environment to test different hypotheses, and it affords assessing a wide range of outcome variables. After reviewing relevant literature and describing the tool and its development, we present a proof-of-concept study that evaluates OGAR's usability and performance and illustrates some ways that it can be used to study the psychology of virtual visits.

Psychological research in art museums

To keep up with patrons, museums need to house objects that are important not just to individuals but to groups of people. Further, they must be able to present those objects in a way that is meaningful to those groups. This has led to recent efforts to recast museums as more "user friendly" and to engage visitors in a more participatory way (Choi, 2013). Increasing emphasis is being placed on identifying who the

✉ Rebekah M. Rodriguez-Boerwinkle
rmrodriguez@uncg.edu

1 Department of Psychology, University of North Carolina at Greensboro, P.O. Box 26170, Greensboro, NC 27402-6170, USA

2 Department of Computer Science, University of North Carolina at Greensboro, Greensboro, NC, USA

audience is, how they interact with individual objects or entire galleries, what information they take home with them, and what meaning they assign to their experiences (Brieber et al., 2015; Leder et al., 2012; Smith, 2014).

Traditionally, the psychology of art has worked to answer some of these questions through studies of individual artworks in lab settings, which offer superior controllability. However, the field recognizes that lab settings do not offer the proper context under which artworks are normally viewed. For example, participants who freely visited an exhibition in a museum viewed the artworks for longer and gave them higher subjective liking and interest ratings than participants who viewed the same exhibit in the lab (Brieber et al., 2014), and participants had greater affective appreciation for, and memory of artworks viewed in a museum context (Brieber et al., 2015). There are increasing efforts to study the psychology of art and aesthetics in real-world, ecologically valid contexts such as museums, galleries, sculpture gardens, and street art sites (Mitschke et al., 2017; Pelowski et al., 2017; Specker et al., 2017). By and large, art museum research uses both recruited and natural visitors for participation. Participants are often given questions or task instructions before beginning their visit and asked to complete some questions or tasks during or after the visit. Researchers often use mobile eye tracking units (Garbutt et al., 2020; Santini et al., 2018), GPS devices (Estrada-Gonzalez et al., 2020), tablets (Cotter et al., 2021; Rodriguez et al., 2021) smart phones (Specker et al., 2020), or simply pens, paper, clipboards and stopwatches (Smith & Smith, 2001, 2006) to capture data during each participant's visit. Data collection may be primarily interested in background measures and post-visit responses, or focused on eye gaze, viewing time, social behaviors, and movement path (Pelowski et al., 2014, 2017).

There are many strengths to this type of field research. Studying art viewing in museums provides richer context and therefore greater ecological validity than lab studies. In addition, the museum context lends itself to stronger aesthetic experiences with greater appreciation and engagement with the artworks (Brieber et al., 2014). Researchers can examine relationships between individual artworks or the entire visit as a whole unit instead of individual works (Smith, 2014). One can also investigate the effects of spatial features like room size or wall color and intentional curatorial choices regarding theme, style, artwork placement, lighting, and accompanying text (Pelowski et al., 2017; Specker et al., 2020). Finally, social interactions with other visitors present in the gallery can necessarily only be studied in a social space (Pelowski et al., 2014).

Unfortunately, field research in museums also has its challenges. It is often as hard to take research into a museum as it is to bring in a bottle of water. Museum staff can be wary of outsiders, so trusting relationships and effective partnerships take time to build. Once researchers are in the building, willing participants can be difficult to attract. All told, most researchers who conduct field research in museums would agree that it is intensive in time, labor, and research personnel.

Another challenge involves manipulating field environments. Few curators will allow researchers to vary aspects of their exhibits (see Reitstätter et al., 2020, for a good example). Other changes, like room size and wall color are simply impossible to alter for the sake of an experiment. And as recent experience shows, data collection may be limited or impossible in times of public health crises and other events that limit access to field sites.

Current virtual art gallery tools

Field studies in museums have many strengths yet pose significant challenges for researchers. One way to balance the trade-off of realism and control is to use virtual gallery tools and simulations. While there's nothing quite like being in a real museum, virtual gallery environments offer an opportunity for a middle ground between the realism of a museum environment and the controlled-but-sterile environment of a research lab.

In recent years, many museums, galleries, and presentation venues have turned to virtual environments for a wide range of uses in addition to their traditional in-person spaces. For example, schools and educational environments may use virtual spaces to provide in-depth exploration and experience-based educational activities. Museums, galleries, and cultural sites may use them to reach people who are not able to visit otherwise or to showcase elements of their collection that are not often on physical display.

The emergence of virtual spaces as innovative and practical alternatives to traditional spaces has been further fueled by several factors. The explosion in virtual reality and growing interest in virtual media have both been big contributors to growing desire for virtual environment tools in business, starting as far back as the 1990s (Leston, 1996; Patel & Cardinali, 1994). The outbreak of the COVID-19 pandemic and resulting rise in social distancing measures aimed at closing access to public spaces has been another (Agostino et al., 2020). Although there have been a wide range of implementations for virtual spaces, here we will review those used primarily for displaying and sharing artwork.

Perhaps the most well-known tool has been the Google Arts and Culture Project. First launched in 2011, Google Arts and Culture has since partnered with over 2000 major museums and cultural institutions around the world to create online simulations of entire museum spaces for free to both the partnering institution and online visitor (Google Arts and Culture Project, 2011; Proctor, 2011). Their process works

by using a trolley system to take thousands of pictures of a museum's interior and digitally stitching them together to create a 3D environment. Then, using software developed for Google Street View, users can navigate the space using a process known as animated interpolation, whereby a person clicks on a point in the distance and undergoes a smooth ("animated") viewpoint transition ("interpolation") from one point in space to the other (Moghadam et al., 2020). This method is somewhat akin to teleportation, but the position change is not instantaneous; instead, avatars are slid along a line from point A to point B.

Another class of popular tools, two of the most popular of which are Artsteps and CAPTURE3D, allow a creator to personalize digitally rendered 3D spaces and share those either privately or publicly. These tools are useful because they allow the user to customize their own virtual spaces and upload their own images using intuitive graphic user interfaces and canned design features. They also allow additional features like over-screen informational pop-ups when artwork is hovered over or clicked on. Movement for these spaces also takes advantage of animated interpolation.

Finally, more intensive tools have been suggested, such as Ikei et al.'s (2013) virtual experience system for digital museums, which uses "a three-dimensional visual display, a spatial sound, a haptic/tactile display for a hand and foot, a wind and scent display, and a vestibular display" (p. 204) to create a multisensory theater aimed at use in interactive exhibits. This type of tool, however, has not achieved wide use.

Using virtual gallery simulations for basic and applied arts research has the potential to overcome many of the challenges associated with traditional museum research. With a little help from online survey platforms, researchers can easily access large, diverse online samples. Virtual spaces are also easily manipulated: several available options allow gallery designers to manipulate floorplans, wall and ceiling colors and textures, and artwork size and placements. Finally, online data collection is safe during public health crises and accessible on most computers, bypassing difficulties in transportation and access.

Unfortunately, there are also limitations with existing tools that constrain their capabilities for research use. First, the process used by Google and digital tools like Artstep and CAPTURE3D are too expensive to be practical for research use. Second, none of the currently available tools are extensible, which prevents researchers from modifying applications to ask new questions. This severely limits their ability to collect and export research data for analysis of how the virtual visitors engage with the environment and artworks, such as how and where they move and what they view. Finally, systems using animated interpolation—although clearly preferred due to its ability to translate to mobile or touchscreen devices—are visually disjointed, which limits

the ecological validity of virtual galleries when used as proxies for in-person experiences and tend to create motion sickness (Moghadam et al., 2020).

The Open Gallery for Arts Research (OGAR)

To provide a low-cost, versatile, and extensible tool for researchers interested in studying the unique characteristics of virtual art gallery spaces that are becoming increasingly common additions to traditional exhibits, or for those seeking greater ecological validity than lab studies but greater control than traditional museum environments, we created OGAR, the Open Gallery for Arts Research. OGAR is best understood via a *see-it-for-yourself* approach, so a sample walk-through video is available for viewing at Open Science Framework (https://osf.io/cwumb/). OGAR is composed of two parts: The OGAR Client, which presents the gallery to the user; and the OGAR Server, which receives and records activity information from instances of the Client. The OGAR Client runs individually on each participant's computer, while the Server runs on an Internet-connected server. Our study integrated the OGAR Client in a page of a Qualtrics survey, but it can be used standalone, implemented in lab-based software, or integrated into most online survey providers.

User interface

From a user perspective, the gallery is experienced as a simple 3D space with a first-person viewing perspective. For this application, we chose to use keyboard-controlled smooth movement with mouse free-look. The user can change where they look by moving their mouse, and they can change their avatar's location by holding the arrow or W, A, S, and D keys on their keyboard. This choice was informed by informal control and interface best practices that have gained popularity in recent decades for applications and games using 3D first person perspectives (Laramee, 2002, as cited in Whitty et al., 2010). Users, via their avatars, move freely throughout the space within the walls of the researcher-designed gallery layout. Movement speed accelerates to a standard walking pace of 1.8 m/s, and artworks are sized to reflect the true size and proportions of those pieces in real life. Due to its wide usage in film and media and documented preference by the viewing public, aspect ratios are set at 16:9 (Nystrom & Fairchild, 1992). The gallery is set to visually refresh at the device's screen refresh rate (typically 60 fps), and resolution is device dependent and varies by participant.

The user can see floor, ceiling, and walls that are colored and flat. In addition, predetermined artworks are clearly visible hanging on the walls of the gallery. The gallery's features can be easily modified by the investigator with little

limitation. For example, researchers can vary the floorplans, the artworks and their placements, the colors of the floor, ceiling, and walls, the gallery lighting, movement controls, and environment physics. As the user interacts with the gallery, their position and view are recorded.

System architecture

To collect participant data, Internet-connected infrastructure is required. Our study used two servers and Qualtrics. The OGAR Client was embedded into a Qualtrics survey using Qualtrics's *Add JavaScript* feature on an otherwise empty question. This embedded JavaScript includes only the Client program but does not include any gallery contents. First, the OGAR Client reads configuration in its environment to determine what it should present. In our study, it used Qualtrics Embedded Data to determine which gallery plan should be presented. Next, the OGAR Client fetches the gallery definition, art images, and other resources from a static file server. This server operates as a typical HTTPS server and can serve the gallery contents publicly over the Internet. While these resources are being retrieved, the client displays a loading screen to the user. As a final preparation step, the Client connects to the OGAR Server and prepares to send interaction data. When these steps are complete, the OGAR

Client displays the gallery to the user. As the user interacts with the Client, it sends position and view information data to the OGAR Server. In addition, other events—such as gaining and losing browser focus and full screen status—are sent to the Server as they occur. A diagram of OGAR system development can be seen in Fig. 1.

Technologies used

The OGAR Client is written in JavaScript and executed within participants' web browsers. It uses the standard WebGL version 1 interface (*Web Graphics* library; Khronos Group, 2011), which is a web standard for the development of web browser compatible 3D graphics interfaces, to render the gallery to the user in an HTML canvas element (Mozilla., 2021). During use, the Client program opens a WebSocket to the OGAR Server and sends updates to record the user's actions. The Client also interacts with the Qualtrics JavaScript API for interacting with Qualtrics Embedded Data and controlling survey flow.

The OGAR Server is a Python3 script that uses the Python *WebSockets* library (Augustin, 2021) to receive connections and data (in this case in-gallery user movements, view direction changes, and other application events) from the Client. The resulting data is recorded in a SQLite3
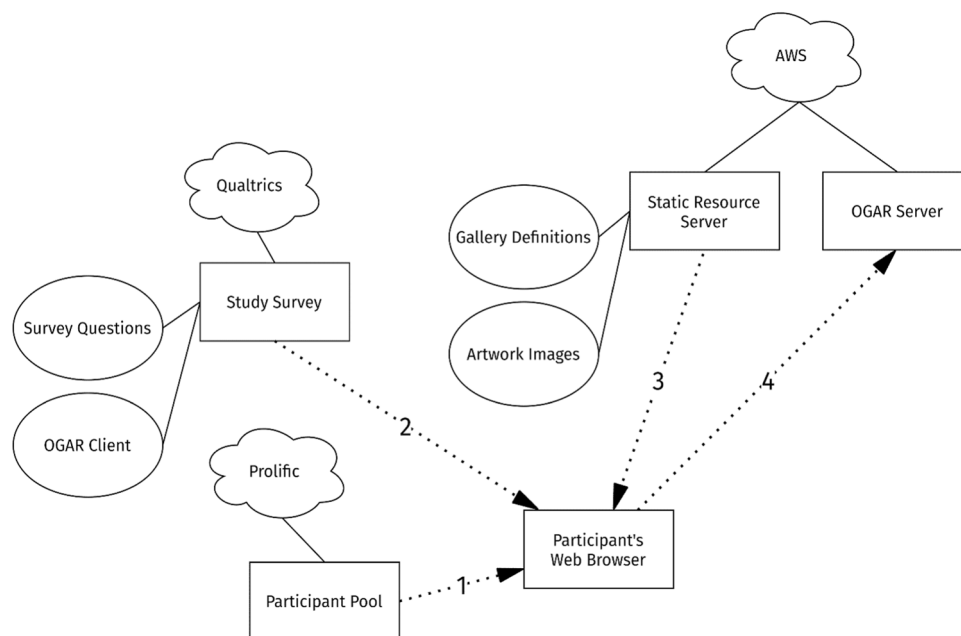

Fig. 1 Diagram of OGAR system deployment. Note. Participants start by being assigned the study through Prolific (1). Next, participants are directed to Qualtrics, where they connect to the survey (2). The survey contains the OGAR client. The OGAR client connects via the Participant's web browser to the static resource server hosted on AWS to retrieve its gallery definition and required art images (3). Finally, the OGAR client connects to the OGAR server to record the participant's actions (4). In this diagram, "clouds" are service providers, "boxes" are semi-tangible architectural elements, and "ellipses" are general resources owned by the boxes. *Solid lines* represent ownership, and *dotted lines* represent the action of data being transferred to and from the participant's web browser as they interact with the overall system

database (Hipp, 2021) where each client connection by a study participant has a random identifier, which allows reconciliation and linking with Qualtrics study results. For this study, we executed the OGAR Server on a Debian 10 server running on an EC2 T3.Micro cloud instance from the cloud provider Amazon Web Services (AWS).

Gallery definition

The OGAR Client displays a gallery based on its *gallery definition*, which is created by the researcher according to their needs. These definitions are formatted as JSON Objects (ECMA-404, 2017). The definition format is minimal, but it is designed to be easily extensible as future studies require. In this study, two premade, static JSON files were used, but this JSON data could potentially have any source, such as existing architectural plans for real-life spaces, or could even be generated based on study data. For a gallery to be defined completely, several parameters must be outlined.

In terms of defining the room features of a virtual space in OGAR, the gallery designer must specify a "walls" list that determines the location of walls in the gallery. This list contains sub-lists that stand in for chains of walls where the end point of one wall is the start of the next wall. Each wall chain is an even-numbered length list where every two numbers determine the Cartesian coordinates of the next wall point. All wall chains have a minimum length of four numbers (two (X,Y) pairs) to define a single wall. For example, a value of "walls": [[0,0, 1,0, 1,1], [5,5, -5,5]] defines two wall chains. The first wall chain consists of two walls—one from (0,0) to (1,0) and the second from (1,0) to (1,1). The second chain consists of a single wall from (5,5) to (-5,5). The whitespace between coordinate pairs is optional within the constraints of the JSON format. The gallery designer must also set "wallHeight," a numerical value that determines the height of walls in meters. A value of 3.05 was used in this study. Finally, a "texture" string must be included to describe the source location of an image which defines the colors of the walls, floor, and ceiling. This image will be retrieved from the static resource server. For the current study, this image consists of three pixels, one for each of the three colors, but it could be extended to instead have three repeated textures.

To place artworks or other stimuli within the virtual environment, the creator must specify an "art" object that contains uniquely named definitions for each artwork in the gallery. Each art definition is an object that must include a "size" list containing the dimensions of the art and a "texture" string that describes the location of the source image on the static resource server. For example, the Mona Lisa may be defined as "LV_MonaLisa":{"size":[0.53, 0.77], "texture":"img/monalisa.jpg"}. The gallery designer must also define an "artPlacement" array of objects. Each object in the array describes the specific placement of an artwork

within the gallery with the following parameters: an "art" string that references a member of the top-level "art" object, a "dir" value that indicates the orientation of the art around the vertical axis in degrees, and a "loc" coordinate array that provides the X and Y position of the artwork in the gallery. In addition, a "height" value must be specified that determines the height of the center of the artwork from the ground.

Finally, to facilitate the first-person experience of the gallery, a "patron" object must be defined to set information related to the user's avatar. This includes a "height" numerical value that determines the user's eye height and a "start" coordinate pair list that sets the user's initial location. For this study, we placed the user's eye height at 1.65 m and specified their start location at [0,0] (the center of the room) within each gallery.

Data format and collection

WebSocket is capable of full-duplex communication, but in this application the communication is unidirectional, and no data is sent to the client from the server. Upon loading, the OGAR Client connects to the OGAR Server running at a preconfigured Internet address. Once a connection is established, the Client sends introductory data. After that, the Client reports the avatar's position within the gallery every 200 ms and other events as they occur. To avoid inaccuracy stemming from variable network delays, or jitter, caused by congestion and other factors, every message is timestamped by the OGAR Client.

Data is recorded by the OGAR Server in a relational database, which has multiple tables connected by reference keys. The primary table, titled *participant*, contains all connections made to the Server by Clients. Each connection is assigned a unique identifier, and Clients may also pass their own self-reported identifiers. In this study, Clients passed an ID generated by Qualtrics, and defined as Embedded Data, as a key for future relational joining with the Qualtrics survey results. This *participant* table also holds assorted other client information, such as connection and disconnection time.

Another table, *position*, records position data for each participant. Each entry in this table is a single position (and view direction), at a single time, for a single user's avatar within the gallery.

Two other tables, *event* and *error*, record events and errors, respectively. The specifics of what is reported may significantly vary in future implementations, but in this study, we recorded events related to mouse-capture in the Client and the full-screen status of the Client. In addition, we created error reports for certain technical problems we thought might arise, but none of those checks triggered during this study.

Data processing

The OGAR Server's collected data goes through several clean-up steps to make it easily ingestible by statistical software. All these steps take place after data collection is finished.

In particular, OGAR reports unprocessed timestamps as either integer UNIX Epoch seconds alone (for errors and events) or in combination with integer milliseconds (for position data). These timestamps are converted into seconds as floating-point values with the time origin at the connection time for the associated Client. Periods when the participant was inactive (as determined by them exiting full screen and surrendering avatar control) were removed from these recomputed timestamps. This allows statistical software to operate purely on when the participant was active as a single contiguous chunk of time.

Another data processing task is view-determination. A custom utility program recreates the gallery for each position table entry (i.e., at each timestamp) and records what the participant in that position was viewing. This calculation determines the first intersection of a ray originating at the avatar's eye and traveling in the direction of the center of their view. The resulting intersections are labeled *wall*, a specific named artwork, or *nothing* depending on what the participant is viewing.

Criteria for data elimination

The OGAR client may not function correctly on every participant's personal computer. Projects using remote samples (e.g., online survey panels) can usually enforce some software or hardware restrictions as eligibility criteria, but many factors, such as software versions and background load, affect performance. Because of this, some participants will create data that should not be considered for analysis. As an example of selection criteria, for the current study participant data were excluded based on the following in-gallery behaviors, indications of abnormal loss of connection, and apparatus-specific signs of poor or malfunctioning browser performance:

- The participant never controlled their avatar with the keyboard. (The avatar's position never changed within the gallery.)
- The participant never moved their mouse. (Their view direction never changed.)
- The maximum distance traveled between avatar position updates was too low. (This is an indicator of poor performance, since position reporting occurs every 200 ms regardless of load, but movement happens uniformly, which may be impacted by excessive load.)

- Events related to mouse-focus and full screen were not reported in rational patterns (e.g., if a client enters full screen, they should exit full screen before the client exits the gallery and continues with the Qualtrics survey). These conditions were likely related to uncommon browsers, behavior-altering browser extensions, or failed browser-mandated user-confirmation checks.

There is significant variability in the performance and functional characteristics of browsers on personal computers, so it is expected that at least a few participants would encounter poor or incorrect functioning, but these measures represent the minimum criteria needed for the gallery to provide a roughly equivalent experience between users.

Cost

In the spirit of accessibility to researchers with a wide range of backgrounds and resources, we designed for resources that are relatively accessible and affordable. As an example, the complete OGAR System set up for use during this study used one AWS T3.Micro instance with 8GiB EBS storage for static resource serving and one for the OGAR Server (~$8/month each). Network bandwidth to and from these two servers was included in the free-tier of AWS, thus incurring no additional cost. In addition, we purchased two domain names and paid $0.99/each/year for 1.111B Class .XYZ domains, but domain name access varies and is provided by some institutions. All told, the entire OGAR System was implemented for less than $20 per month of data collection for the current study. Setting up an AWS server to run with Qualtrics and recruiting paid Prolific participants for participation in our study proved to be a cost- and time-effective approach for our team, but OGAR can be set up to work in a variety of formats. For example, OGAR could be imbedded in a free online survey software instead of Qualtrics, or given developmental changes, in lab-based software so that data collection could be done with student or community samples without online tools, and AWS, of course, could be exchanged with a number of other server set-ups.

Evaluating OGAR

In the present research, we collected "proof of concept" data to assess the potential of OGAR as a tool for studying visitors within a virtual art gallery. A sample of adults was recruited from an online research participant panel (Prolific.co), and the participants were allowed to freely explore the virtual gallery and view the artworks within it. We focused on the OGAR's performance in two key areas: gallery usability and measurement validity. Gallery usability was evaluated using participant responses on the

System Usability Scale (SUS; Brooke, 1996), self-reported nausea, and open-ended reports on user experience immediately after exiting OGAR. The usability data were collected to inform the participants' experience of navigating and interacting with the gallery and to discern how "user friendly" they found it.

Measurement validity was evaluated by manipulating aspects of the gallery and measuring behavior within it. We focused on some fundamental hypotheses that, while obvious and perhaps banal, would nevertheless have to be true for researchers to have any confidence in the validity of OGAR as a research tool. For validity data, we manipulated the size of the gallery—one room or two rooms—as a between-person variable. The two-room gallery had double the number of artworks and double the area, so the manipulation afforded testing some critical assumptions of successful use: (1) as the gallery space increases, participants will spend more time within it; and (2) as the gallery space increases, participants will travel a greater distance when navigating it.

Finally, for further evidence for the measurement validity of OGAR, we evaluated whether participants interacted with the artworks—that is, whether their time and movement within the virtual gallery was guided by the artworks as opposed to random or listless movement. Participants' positions in the gallery, movement trajectories, and viewing points were analyzed to discern how they traveled through the gallery, where they stopped, and what they viewed. Taken together, the usability data and the participants' behavior within the gallery should shed light on the value of OGAR as a tool for research on virtual art spaces.

Method

Participants

The present study was approved by the University of North Carolina at Greensboro Institutional Review Board (Study #21-0311), and all participants provided informed consent. A total of 61 adult participants were recruited from the Prolific.co survey panel and paid USD $4.00 for their participation. To be eligible, participants were required to be within the ages of 18 to 70, to be native speakers of English, and to have a minimum Prolific.co study approval rate of 90%. The study was advertised as "desktops only" within the Prolific system (i.e., tablets and smartphones were not permitted, but laptops were). After screening for inattentive responding, drop-out, and technology issues (described in detail later), the final sample consisted of 44 participants—19 women, 25 men—who ranged in age from 19 to 60 (M age = 31.73).

Procedure

Prolific participants were redirected to a Qualtrics survey for the duration of this experiment. People were prompted to provide basic demographic information—their age, country of residence, and gender—before proceeding to the gallery. When the participant arrived at the specified "question," a preview window of the gallery was shown that expanded into full screen when the user clicked on the window. At this point, full controls were enabled, and the participant could navigate the gallery using their keyboard to move cardinally to the view direction. The user could change their view direction by moving their mouse. Participants could peruse the space for as long as they wished. After participants completed their visit, they were able to release their controls, exit full screen mode, and return to the Qualtrics survey by pressing the *Escape* key. The remaining part of the survey involved a series of follow-up questions about their experience.

Artworks Sixteen artworks were selected for use in OGAR, based on prior approaches to artwork selection in similar studies (Belke et al., 2010; Leder et al., 2012). We procured high resolution images from the ARTSTOR digital library and public domain images from WikiArt. A full list of artworks is available in the Appendix Table 4. Where possible, artwork choices reflect those directly used in Belke et al. (2010). However, due to high quality requirements of our application and licensing constraints, some images were replaced with similar works from the same artist or other works. As a rough guideline, we aimed for artwork images between 20 and 50 dpi to ensure high enough image resolution without excess strain on client image download speeds. Artworks were categorized as either representational or nonrepresentational, with equal numbers of each mixed throughout the gallery. The artwork was placed to mimic realistic curation in physical gallery spaces, using aesthetic design principles outlined in Adrian George's *The Curator's Handbook* (George, 2015).

Gallery manipulation Gallery area was manipulated between-person. Participants were randomly assigned to be placed in either a one-room or a two-room version of OGAR. The two-room version appended the additional room directly adjacent to the first room, accessible by an open doorway. The one-room manipulation was enclosed by four walls. Rooms were identical dimensions (10 × 10 m), with the first room of both versions containing the same eight artwork placements and the second room of the two-room version containing an additional eight artworks. Total gallery area and number of artworks were doubled, so that artwork placement in the first room is consistent (with the exception of slightly wider placement between two artworks

to accommodate the doorway in the two-room version) for both conditions, and comparisons concerning number of artworks and distances are facile.

Measures and outcomes

Browser data Qualtrics was set to capture each participant's browser type, browser version, operating system, screen resolution and user agent. This information was used to investigate poor gallery performance in specific cases, so that the system can be improved in later study iterations.

Gallery data The gallery receiver server collects time-based position and gaze data for each participant every 200 ms. Location is recorded in X and Y coordinates with one unit corresponding to one meter of distance in the gallery. Gaze data consists of yaw and pitch and is defined in terms of radians.

User feedback Usability for OGAR was qualitatively assessed via user feedback from the SUS, as well as a few additional questions specific to the gallery, a directed-response item to flag inattentive responding (Maniaci & Rogge, 2014), and an open-ended prompt for additional comments (see Table 1). Since its initial publication, the SUS has been widely used in human-computer interactions research and product evaluation for computer systems (Lewis, 2018). The SUS assesses perceived usability through a 10-item questionnaire with response options scaled from 1 (*strongly disagree*) to 5 (*strongly agree*; Brooke, 1996), and it is designed to be implemented following task-based usability testing. Items are all first-person statements about personal user experience, like "I thought the system was easy to use" and "I found the system unnecessarily complex." In

the present study, the word "system" was replaced with the more specific descriptor "virtual gallery" in line with wording recommendations put forth by Lewis and Sauro (2009).

To create an overall score from the 10-item SUS, all participant responses are shifted so that the lowest possible score for each item is 0 and the highest possible score is 4. Then, even-numbered items are summed and odd items are each subtracted from the sum of the positive scores. The resulting total is multiplied by 2.5, which converts the range of possible values from 0 to 100. A score of 80 is commonly used as a threshold for good system usability (Lewis, 2018). Internal consistency measures for the SUS range from $\alpha = .83$ to $\alpha = .97$, with most studies placing it at about $\alpha = .90$ (Lewis, 2018).

In addition, the two questions explicitly about navigation and art viewing in the gallery were presented with the SUS but treated as separate, individual items during analysis (see Table 1). Participants were also asked what type of input device they used in the gallery (possible responses included *mouse*, *touchpad*, *touchscreen*, *trackpoint*, or *other*), and to report feelings of nausea, they responded, using a 1 (*No, not at all*) to 7 (*Yes, very strongly*) scale, to "Did you feel motion sick, dizzy, or nauseous from the virtual gallery?". Finally, participants were invited to leave open-ended feedback or comments regarding their experience.

Results

Data processing and reduction

Data processing and statistical analyses were conducted in R 4.1 (R Core Team, 2021). Out of the 61 participants who

---

Table 1 Usability questions

System Usability Scale (SUS): 10 Items

    1. I think that I would like to use this virtual gallery frequently.

    2. I found the virtual gallery unnecessarily complex.

    3. I thought the virtual gallery was easy to use.

    4. I think that I would need the support of a technical person to be able to use this virtual gallery.

    5. I found the various functions in this virtual gallery were well integrated.

    6. I thought there was too much inconsistency in this virtual gallery.

    7. I would imagine that most people would learn to use this virtual gallery very quickly.

    8. I found the virtual gallery very awkward to use.

    9. I felt very confident using the virtual gallery.

    10. I needed to learn a lot of things before I could get going with this virtual gallery.

Additional Study-Specific Items

    I was able to clearly view all the artworks present in this virtual gallery.

    I was able to easily navigate through this virtual gallery.

Items were scored on a 5-point scale (1 = *strongly disagree*, 5 = *strongly agree*). The items were presented in a random order

began the study, 4 participants dropped out mid-study and didn't complete the entire Qualtrics survey, and their data were excluded from analysis. Participants were also excluded if they failed a directed response item embedded in the gallery usability survey ($n = 3$ excluded for this reason). These eliminations left 54 participants who were then processed for gallery performance quality. After careless in-gallery behaviors, indications of abnormal loss of connection, and poor browser performance were assessed, we were left with a final sample of 44 participants from ten different countries. The ten participants who were dropped during processing for performance quality can be broken down further: one person experienced total gallery failure with no known cause; one person was dropped for being unable to control their gaze due to using a nonstandard input device instead of a mouse (this participant clicked "other" when asked about their input device and had no recorded movements in their gaze data); and eight people were eliminated for slow movement speed (there are various reasons, from browser-specific issues, to high nausea, why this may have occurred).

Once data processing was complete, analysis was conducted using the R packages *psych* (Revelle, 2021), *reghelper* (Hughes, 2021), and *parameters* (Lüdecke et al., 2020). Gender responses were coded as binary (female = 1, male = 0). In addition, mouse input devices were recoded as binary (mouse = 1, all other input devices = 0) to better reflect our choice to design the gallery explicitly for mouse usage. Nausea, SUS scores, maximum movement speed, total visit and artwork viewing times, and distance traveled within OGAR were explored in the Pearson's *r* effect size metric, using guidelines of .10/.30/.50 to represent small, medium,

and large effect sizes respectively (Cumming, 2012). For categorical participant factors like gender and whether they were using a mouse as their input device, we used Cohen's *d*, which can be interpreted in terms of small, medium, and large effects using .20/.50/.80 as common benchmarks (Cumming, 2012).

Usability

We started by evaluating OGAR's usability through feedback on the SUS and accompanying measures. The SUS had high internal consistency reliability (Cronbach's $\alpha = 0.89$) that was in line with previous work using the scale (Lewis & Sauro, 2009). On average, participants gave OGAR a good SUS rating (Mdn = 87.50 out of 100, M = 82.90, SD = 14.64, range from 37.50 to 100). Both the median and mean were higher than the common benchmark score of 80 used to mark good system usability (Lewis, 2018).

To provide a more granular view of participants' usability experience, Fig. 2 displays a ridgeline plot of the ratings for all 10 SUS items (on their original 1–5 response scale used by the participants). The item-level distributions show that, for seven of the ten items, the modal rating reflected the highest usability option.

To supplement the classic SUS questions, we asked participants whether they were able to clearly view the artworks present in the gallery (Mdn = 5.00, M = 4.57, SD = .79) and easily navigate through the virtual gallery (Mdn = 5.00, M = 4.55, SD = .76). The high scores at the ceiling of the response scale suggest good usability for these specific aspects of the gallery.
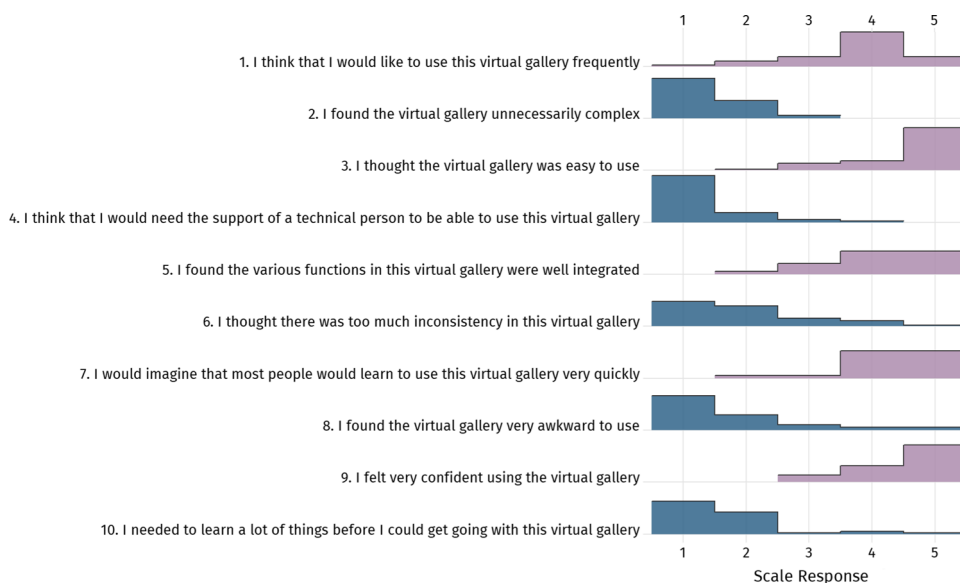

Fig. 2 SUS item score distributions. Note. The figure displays the scores in their original 1–5 response scale (i.e., prior to reverse-scoring and rescaling)

Usability ratings were high on average but exploring variability in usability ratings can give insight into likely predictors of usability experiences. One particularly interesting factor is the experience of nausea. As Fig. 3 shows, nausea ratings were very low, and notable nausea occurred in only a small portion of our sample (only four participants provided nausea ratings of four or greater out of 7; M = 1.36, SD = .97). Ratings of nausea had a modest correlation with SUS scores ($r = -.23$ [$-.49$, .07], $p = .136$), reflecting lower usability ratings as nausea increased. We suspected that poor gallery functioning may have contributed to the nausea experienced by some participants, so we examined whether there was a correlation between nausea and maximum movement speed as a proxy of overall gallery functioning; no such relationship was found ($r = .09$ [$-.21$, .38], $p = .559$).

Because OGAR was designed for use with a mouse in mind but data collection for the current study depended on the personal equipment of our online participant pool, the relationship between input device and usability is important to consider. Participants who used a traditional mouse in lieu of other alternatives gave non-significantly higher overall SUS ratings than those who did not ($d = .24$ [$-.40$, .86]). The SUS items regarding desire to use the virtual gallery ($d = .61$ [$-.06$, 1.26]), finding the gallery unnecessarily complex ($d = -.40$ [-1.03, .25]), feeling confident using the virtual gallery ($d = .37$ [$-.27$, 1.00]), and needing to learn a lot of things before getting going with the gallery ($d = .52$ [$-.14$, 1.17]) conveyed relatively stronger effect sizes for mouse usage.
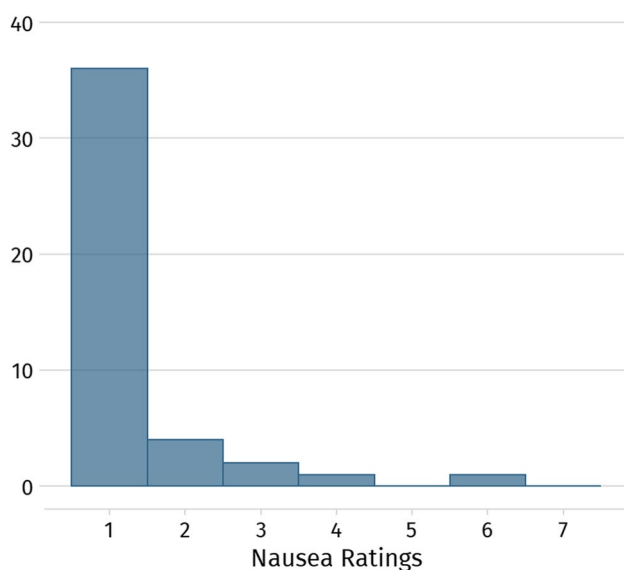
These average scores on the SUS and additional usability questions represent the bulk of user experiences. Most user feedback was positive—something that is reflected in open ended feedback. Many participants wrote that they enjoyed their experience, "nearly felt like [they] were there," and that OGAR was "the easiest [virtual space] to use that [they've] encountered so far." Some participants also provided commentary about their subjective experiences with the artworks: "It was great to see some abstract paintings and some of them were really made me think a lot." Collecting open-ended feedback from our participants also allowed us to hear any specific problems they encountered and additions or changes to the gallery that they would be interested in seeing in the future. For example, one participant's comment that "the art closer to the right of the screen were harder to see and navigate to" within the square gallery condition may imply that the artwork on the right wall, relative to the starting location, may have been too small for adequate viewing on smaller screens by a diverse audience. We also learned that some participants would prefer navigation and exiting instructions available after entering full-screen mode, or that other participants are interested in the ability for in-gallery behaviors that mimic videogames (e.g., a sprint mode) or other applications they often use. All comments can be viewed on OSF (https://osf.io/f9e8d/).

Behavior in the virtual gallery

Following our second aim—appraising the validity of OGAR as a research tool—the position and gaze data collected within OGAR allowed us to identify whether patterns in participant behavior align with expected behavior in physical spaces. Linear regression models were used to examine predictors of participant behavior; the reported effects are standardized (β). For comparisons using categorical predictor variables, such as room condition (one room = 1, two rooms = 2) and mouse use (did not use mouse = 0, used mouse = 1), and continuous outcomes, we reported Y-standardized regression coefficients, noted as β_Y, in which only the outcome variable is standardized (Long, 1997, chap. 2). The coefficients of these regressions are equivalent to Cohen's d effect sizes or the difference, in SD units, in the outcome between both groups (Long, 1997). Descriptive statistics for each room condition can be found in Table 2.

Visit duration On average, people spent about 76 s in the one-room condition and 174 s in the two-room virtual gallery (see Table 2). Thus, in line with our core hypotheses about validity, time spent in OGAR was significantly greater for the two-room gallery condition than the one-room condition ($\beta_Y = .67$ [.08, 1.25], $p = .026$). Time spent in the gallery was not significantly related to nausea severity ($\beta = .11$ [$-.20$, .42], $p = .465$) or to SUS scores ($\beta = .06$ [$-.26$,

Fig. 3 Distribution of nausea ratings. Note. The figure displays participant ratings for the item "Did you feel motion sick, dizzy, or nauseous from the virtual gallery?" on a scale from 1 (*No, not at all*) to 7 (*Yes, very strongly*)

Table 2 Descriptive statistics for each room condition

	One Room		Two Room	
	M (SD)	Mdn	M (SD)	Mdn
Visit duration (*s*)	76.32 (55.44)	63	174.16 (178.23)	116
Distance traveled (*m*)	35.79 (23.05)	36.43	102.48 (48.23)	88.62
SUS total score	83.29 (16.63)	90	82.60 (13.28)	87.50

Participants were randomly assigned to the one room (*n* = 19 [nine women, ten men]) or the two room (*n* = 25 [ten women, 15 men]) condition

.37], *p* = .719). People who used a mouse spent slightly less time in the gallery, but not significantly so ($\beta_Y = -.33$ [– .97, .32], *p* = .311). In sum, visit length was greater when OGAR presented more rooms, and comfort and usability had non-significant relationships with the time that people chose to spend in the gallery.

Distance traveled What factors affected the distance people traveled within the gallery? As expected, participants traveled further in the two-room condition than in the one-room condition ($\beta_Y = 1.30$ [.83, 1.77], *p* < .001), supporting one of our core hypotheses about validity. In addition, people who spent more time in OGAR traveled a further distance within the gallery ($\beta = .69$ [.46, .92], *p* < .001). Finally, total distance traveled within the gallery was only weakly and non-significantly related to nausea level ($\beta = .15$ [– .15, .46], *p* = .320), overall SUS score ($\beta = -.09$ [– .40, .22], *p* = .540), or mouse use ($\beta_Y = .14$ [– .51, .79], *p* = .661).

Engagement with the artworks Our third aspect of validity—whether people actually approached and engaged with the artworks—was examined descriptively using heatmaps overlaid with regions of interest relevant to each artwork. Heatmap density was calculated via time-stamped *X* and *Y* position data for each participant as they explored the gallery and was weighted evenly for each participant. This ensures that every participant contributed evenly to the heat map density. In addition, density at the starting location for entering the gallery was omitted to prevent any visible heat spiking that is irrelevant to deliberate participant movement. Finally, the heatmap underwent histogram equalization to optimize the global contrast of our data and enhance the level of visible detail in our mapping. Regions of interest were defined by partitioning the floorspace of the gallery into Voronoi cells that comprise a larger diagram (Voronoï, 1908). Each cell represents the region of the gallery that is closer to the center of that cell's artwork than to any other. Once the Voronoi diagram is overlaid on the heatmap, any intense clusters of participant movement should be visible within a specifiable artwork region. Note that this exploratory data visualization method does not yield any inferential

statistical tests, but because it is data-driven, it is robust and fully reproducible.

Figure 4 illustrates the resulting heatmap with overlaid Voronoi regions. Artworks (to scale) with black points at the center of each image are placed on the walls for reference. Areas of red are the "hottest," representing places where the participants spent the most time. Areas of the highest density have additionally been outlined in black for visual clarity.

The diagrams for both room conditions clearly reveal "hot spots" clustered in front of the artwork's center that are most often within the Voronoi region defined by each artwork's location. This indicates that participants' movement within the gallery is purposeful and consistently guided by the artworks, as it ought to be. Additional Voronoi regions with sporadic hot spots can be seen surrounding the center of each room and can be thought of as highly trafficked movement areas or common pathways around the gallery as opposed to destinations of interest.

Illustrating some options and opportunities for researchers

As noted in the Introduction, several available virtual gallery programs have different useful characteristics but have not been coalesced into a tool ideal for research use. Extending our discussion of OGAR past its usability and basic features seems helpful to show what researchers can do with the virtual gallery. These remaining findings are intended to demonstrate some functionality that might spark ideas and give food for thought for researchers interested in using OGAR.

Viewing time Viewing time—how long people spend looking at an artwork—is a major outcome in art and aesthetics research (Carbon, 2017; Pelowski et al., 2017). Studies of free-viewing behavior in museums commonly show that visitors spend much less time viewing an image on a wall than many would think—often between 8 and 20 s (Reitstätter et al., 2020; Smith et al., 2017; Smith & Smith, 2001)—in light of how impactful people later describe the experience (Smith, 2014).

Viewing time is easy to obtain from OGAR. Since every movement and gaze that the participant completes within OGAR is recorded, we can take advantage of existing gallery infrastructure to automatically code what artwork a participant is examining at any given point in their visit in a low-level viewing analysis. To do this, we created a parallel program for view determination that operates on a viewpoint, defined by the set eye height and avatar location within the gallery, and gallery definition (see Fig. 5). To figure out what a ray extending from that viewpoint would hit first (i.e., what a person is "viewing"), every artwork and wall segment are turned into two triangles each, forming a rectangle. Then, a
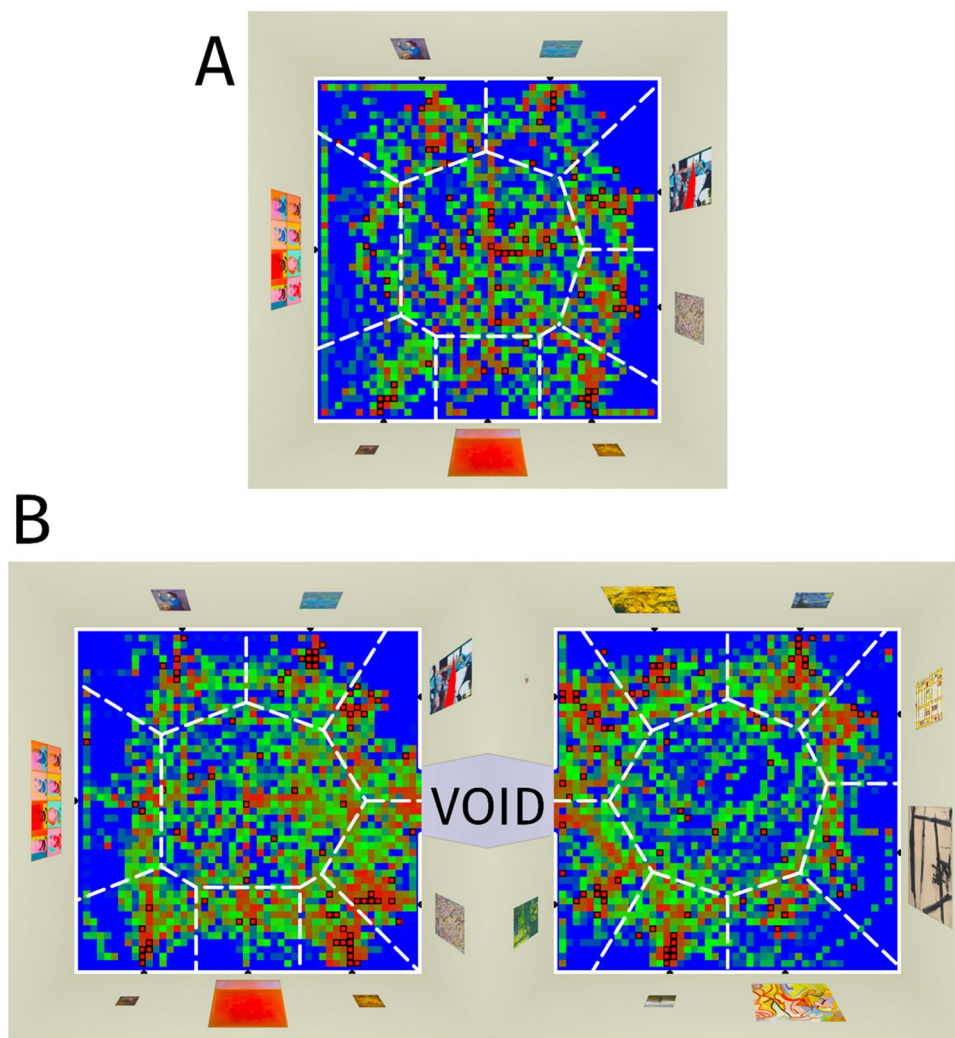
Fig. 4 Heatmaps of one-room and two-room conditions with Voronoi region overlays. Note. The area marked VOID on Fig. 4b represents the doorway between rooms in the two-room condition. No hall-way or area exists at this designation—it's a result of the converging bird's-eye viewpoints

Möller–Trumbore intersection (Möller & Trumbore, 1997) is applied between every triangle and a line defined by the viewpoint. The shortest distance intersection is kept as the view target. If no triangle intersects, the view determination is "None." View behavior can be coded as a binary *yes* (1) or *no* (0) for viewing artwork or categorically assigned with the corresponding artwork, given that participants are viewing an artwork at a given timestamp.

As an example, Table 3 lists the average viewing times for each artwork in the two-room version of the OGAR gallery used in the current study. (We focus on the two-room condition because it has the largest number of artworks.) Overall, gallery visitors in this condition viewed an artwork for a mean of 5.92 (SD = 2.40) s, which falls on the lower end relative to research on artwork viewing time in real-life museum environments. More broadly, people in the two-room condition spent a little over half their time looking at

artworks (M = 94.74 s) as opposed to other features of the space (i.e., walls, or nothing; M = 80.64 s).

Viewing distance Another common measure of interest to museum researchers is viewing distance: how far away, in meters, visitors stand from a work when viewing it (e.g., Carbon, 2017; Clarke et al., 1984; Estrada-Gonzalez et al., 2020). Perhaps unsurprisingly, research conducted in unconstrained field settings commonly finds that viewing distance increases as the artwork size increases. In OGAR, viewing distance in meters can be measured by taking the coordinate location of each avatar at each timestamp that a participant is viewing an artwork and calculating the distance between the location coordinate and the artwork coordinate. Then, viewing distance measurements can be averaged for each participant and the entire sample for each artwork present (see Table 3). To draw once again from the two-room condition

Fig. 5 Multiview representation of view determination. Note. Three ▶ views of an illustrative scene involving a user avatar (*stick figure*) viewing an artwork (*red*). The *dotted line* emanating from the avatar's head indicates the direction that the user is looking in the gallery. In this scene, the ray drafted from the avatar's head is tested for intersections against triangles that compose the walls and artworks. The center of the view ray intersects with the upper-right triangle composing the red artwork. Therefore, this hypothetical user, at this point in time, is determined to be viewing the red artwork. Panels A and B show this interaction from two third-person perspectives. Panel C represents the projection of this scene as a "bird's eye view," which makes the intersection with the artwork more readily apparent

of the present study for an example, participants viewed artworks at an average of 2.04 m (SD = 1.26), although viewing distance varied considerably by artwork (range 0.71 to 5.81 meters).

This picture-to-picture variation in viewing distance, it turns out, is a function of image size. In the virtual gallery, viewing distance was strongly correlated with artwork area ($r = .90$ [.73, .96], $p < .001$). As Fig. 6 depicts, people viewed larger artworks from farther away and smaller artworks from close up, just as visitors typically do in real-world galleries (Carbon, 2017; Estrada-Gonzalez et al., 2020).

Navigation and movement trajectories In addition to viewing behaviors, participant navigation is a common outcome in field studies of museum visits (Tinio & Specker, 2020; Tröndle, 2014): the paths people take as they move through a gallery is interesting in its own right but also practical knowledge for curators and museum professionals. Within OGAR, researchers can similarly explore how people navigate and interact with virtual gallery spaces. Using the participants' coordinates across time, researchers can identify the temporal qualities of movement in the virtual gallery.

For example, Fig. 7 displays the movement trajectories of three representative participants who were randomly assigned to the two-room condition (top three panels) as well as a combined overlay (bottom panel). Although all participants started at the same position, they took different routes through the gallery, explored different rooms first, covered varying amounts of ground, exited the gallery at different spots, and showed differences in trajectory features like the straightness of their path. Researchers interested in movement and trajectory analysis could find the data provided by OGAR fertile.

Going beyond a static snapshot of a participant's movement, we can animate the path a person takes around the gallery. This provides in-depth temporal information for a single person and is an intuitive, holistic way of presenting dense position and viewing data. As an example, Fig. 8 links to an animated video of a single participant's time spent in the virtual gallery. The red line traces their movement; the small green line indicates their gaze direction.
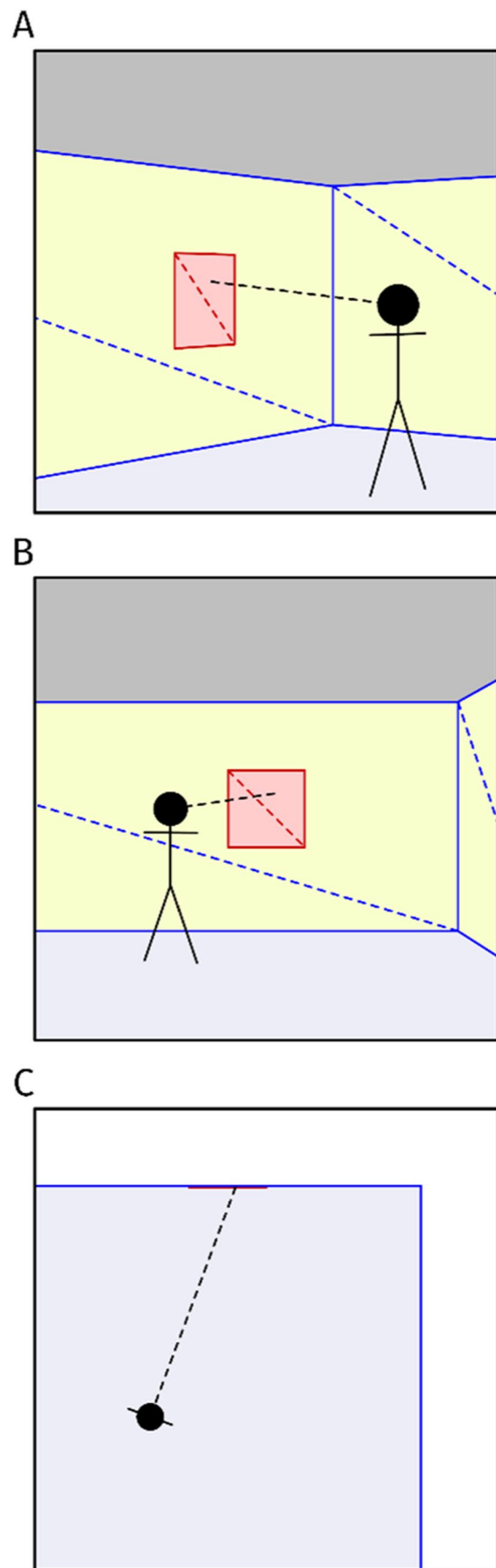
Table 3 Viewing behaviors for the two-room condition

	Viewing time (s)		Viewing distance (m)	
	M	Mdn	M	Mdn
Non-Art Gallery Elements				
Viewing Nothing	11.46	5.60	NA	NA
Viewing Wall	69.18	47.60	NA	NA
Art Images				
Lita Curtain Star	5.89	3.60	3.17	2.44
Woman with a Fan	5.04	3.20	1.55	1.14
Water Lilies	5.07	2.00	1.21	0.97
Untitled (Green and Red; *FR*)	5.62	2.80	2.73	2.17
Terrano X	12.70	4.40	1.57	1.29
Starry Night	6.10	1.40	1.13	0.88
Broadway Boogie Woogie	4.33	3.40	2.81	1.82
Painting Number 2	9.66	5.00	5.81	3.11
Untitled (*WK*)	5.61	2.60	2.41	1.89
Solitary Tree	6.90	3.40	0.98	0.71
Untitled (*PK*)	4.38	1.40	1.56	1.37
Eyes in the Heat	6.73	3.60	2.18	1.54
Reclining Girl	4.22	2.80	1.05	0.98
Untitled (*MR*)	4.27	2.40	2.64	2.35
The Silver Goblet	5.86	2.00	1.14	0.67
Hare	2.36	0.40	0.71	0.49

Details about the artworks are in the Appendix Table 4. Viewing times are reported in seconds; viewing distances are reported in meters. Untitled works are followed by artist initials in italics for ease of identification
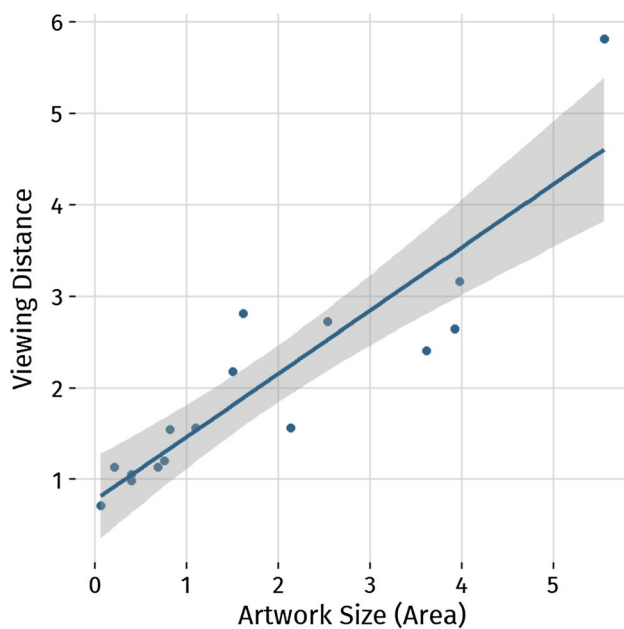
Fig. 6 Relationship between artwork size and viewing distance
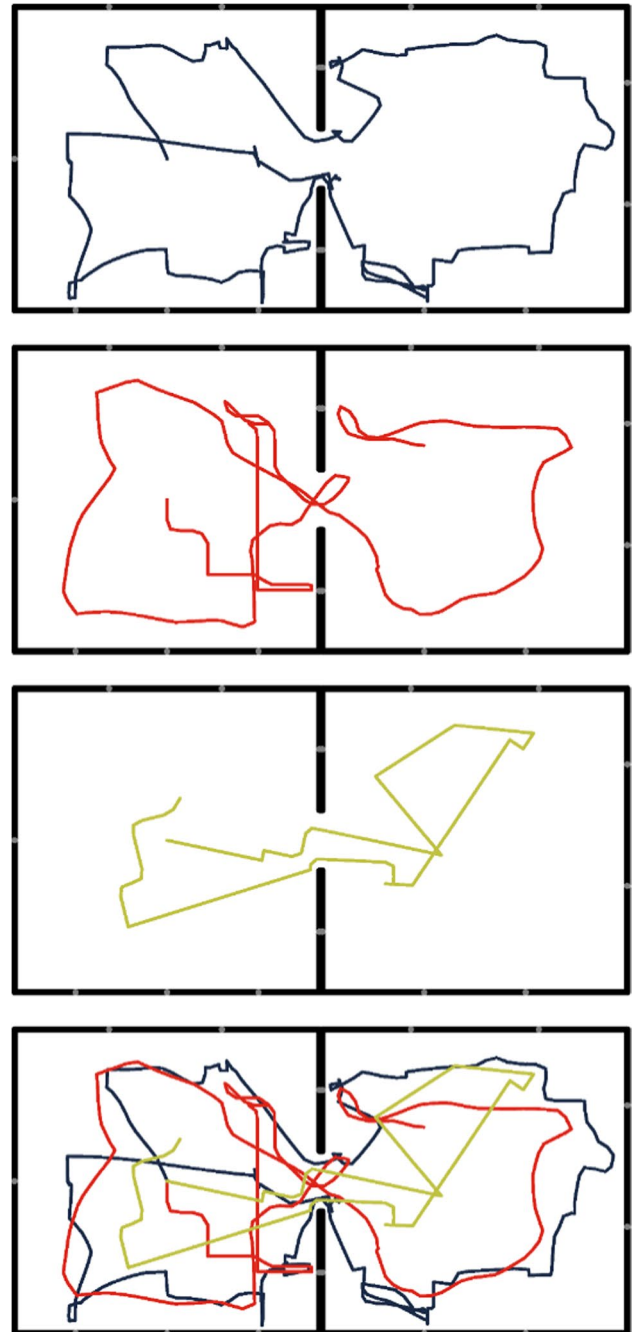
Fig. 7 Individual navigation trajectories for three participants in the two-room gallery (top three panels) and the overlaid patterns (*bottom panel*)

Discussion

In the present research, we developed and evaluated the Open Gallery for Arts Research, or OGAR, as a tool for exploring the psychology of virtual gallery encounters. In contrast to the current landscape of offerings, OGAR is an
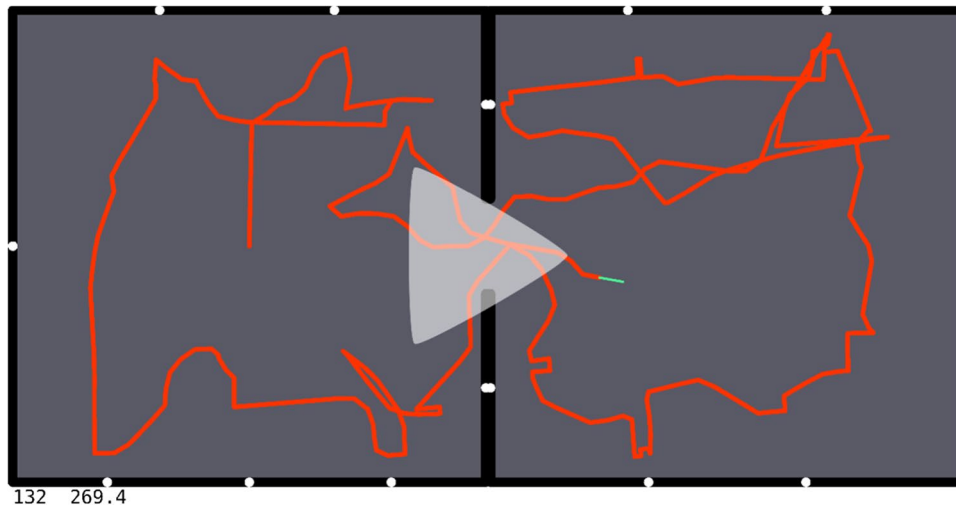
Fig. 8 An animation of a single participant's time spent in the virtual gallery. Note. The image links to a video stored on Open Science Framework (https://osf.io/63rsz/). The red line indicates movement; the small green line indicates gaze direction. The animation is played at 3x speed

affordable, flexible, and extensible open-source tool for creating virtual art gallery spaces and measuring participants' behaviors within it. A proof-of-concept study was conducted to assess the usability and performance of OGAR in an online sample of adults.

First, the usability of OGAR appears to be strong based on results from the SUS, additional gallery-specific usability questions, nausea ratings, and open-ended feedback. Average SUS ratings were high (Mdn = 87.50 out of 100), beyond the threshold of 80 commonly used to indicate good system usability (Lewis, 2018). Variance in SUS scores were related in coherent ways to other factors. The small portion of the sample with elevated nausea ratings gave lower SUS ratings, and using an input device other than a mouse, the system's optimal input, was likewise associated with lower SUS ratings.

Second, the behavior of participants within the virtual gallery was coherent and predictable, resembling what researchers observe in participants navigating real-world gallery spaces. Using the position and gaze data collected within OGAR, we were able to support the view that our online participants were interacting with the virtual gallery in the ways that researchers in the psychology of museum experiences would expect. People who were randomly assigned to a gallery that was twice as large and contained twice as many artworks, for instance, spent a much longer amount of time in the virtual gallery and traveled a much greater distance. While not shocking, such findings show that participants were interacting with the gallery as one would expect. In addition, as evidence that participants used the gallery to view the artworks, heatmaps of the gallery floorplan partitioned into Voronoi regions for each artwork clearly show high densities of participant movement clustered in front of each artwork along with commonly trafficked paths between artworks. These key findings suggest that the OGAR system produces basic participant behavior that is psychologically coherent and similar to gallery behavior in traditional in-person settings (Tinio et al., 2015).

Finally, we sought to illustrate how OGAR can be applied and extended for future research use. We showed how a participant's movement trajectory through the gallery can be identified and visualized, which could be useful for researchers interested in how environmental and curatorial factors influence how people move through gallery spaces (Bourdeau & Chebat, 2001; Tröndle, 2014). In addition, we showed how viewing data can be used to obtain measurements of viewing time and viewing distance, two outcomes of long-standing interest to researchers studying how people view art in museums (Carbon, 2017; Estrada-Gonzalez et al., 2020).

Extensions and options

OGAR is a versatile tool that affords a wide range of opportunities. Researchers can extend OGAR or alter its configuration to fit their specific needs by varying any of the following:

- Gallery layout (i.e., size and configuration of gallery walls)
- Artworks (i.e., image choices, sizes, placement)
- Aesthetics (i.e., floor, ceiling, and wall colors)
- Avatar characteristics (eye height, acceleration, maximum speed).

In addition, OGAR's licensing allows researchers to make more extensive changes to OGAR's software if they wish. Doing so opens the possibility for additional features like audio, in-gallery pop ups, randomization features, or any number of add-ons that a researcher may desire for their work. Changes and additions to the OGAR software can be shared with GitHub pull requests. Updates to OGAR and further details are available on GitHub at https://github.com/mboerwinkle/OGAR .

Behaviors such as artwork viewing time and viewing distance can be recreated using avatar height, gallery layout specifications, and participant movements collected during data collection. These measurements can then be analyzed in relation to researcher-set design features of the gallery like artwork choice, curation, or layout of the virtual space. They can also be examined alongside additional surveys or other measurement tools that can easily accompany OGAR in platforms like Qualtrics. This particularly opens up the possibility of deeper examination of subjective experiences as opposed to the behavioral measures focused on in the current paper. Further, data can be animated to show navigation trajectories in OGAR that can be analyzed qualitatively, examined in terms of artwork regions defined by Voronoi cells, or simply examined between participants. Also of interest, OGAR output may serve as a suitable proxy for mobile eye tracking. Although bounded by the edges of a monitor, unconstrained position and gaze movement within the environment allow participants a high degree of visual exploration during their visit.

Some practical issues

A common problem with many virtual environments, videogames, simulations, or other applications using a first-person viewpoint in 3D environments is visually-induced motion sickness (Kennedy et al., 2000; Keshavarz & Hecht, 2012; Stoffregen et al., 2008). To guard against nausea or motion-sickness-prone participants in the present study, we provided brief warnings in the study's Prolific recruitment ad and consent form. In addition, we asked people to exit the virtual space should they feel dizzy, nauseous, or motion sick during participation. Nausea ratings were quite low in our study, but because these represent the scores of only those people who completed the study to that point and not those who dropped out or who declined to take part due to likely nausea, our data probably underestimates the base rates of nausea experiences in OGAR. We recommend including warnings about motion sickness during participant screening as well as measuring ratings of nausea experienced during participation, which are useful for analyses of participant behavior and for possible exclusions. Further, because some motion sickness is inevitable for studies employing virtual galleries and similar tools, these precautions are important for both ethical treatment of participants and overall data quality.

As with all online tools, the OGAR Client has issues to be addressed related to compatibility between different participants' computer environments. Incompatibility can occur for many reasons, but non-standard web browsers (e.g., outdated, poorly configured, or simply non-compliant) are a major source. In addition, old, underpowered, or otherwise overloaded computer systems can contribute to poor behavior, as with any system that relies on real-time input. Although it is desirable for all participants in online studies to have similar experiences, in practice there is no way to ensure a perfectly identical experience for everyone when research is conducted on personal machines. As such, the best a team of researchers can do is to carefully weigh the values of control and flexibility for a particular aim. For this study, we chose to control hardware and software by dictating that participants must use a desktop or laptop computer with a non-Safari browser. We did not, however, mandate any more stringent hardware requirements like amount of RAM needed, screen resolution, or graphics processor attributes, or require that participants download or have access to specialized software. These initial specifications simply sought to eliminate clearly incompatible participants.

After data collection was complete, a second line of standards was used to determine what level of performance would be considered acceptable. Thresholds for performance based on mouse movements, maximum speed, and event reporting were established to eliminate some participants post data collection. Again, although some level of performance is required for useful data collection, it is not necessary to eliminate every participant who possibly was on the edge of compatibility, and the least strenuous thresholds that are acceptable should be placed to avoid over-filtering the data. Mechanisms for measuring software performance for the current study are discussed further in the introduction, but future iterations of OGAR will likely improve on these by adjusting minimum speed requirements and recording participant frame rate. Ultimately, however, many of these concerns can be sidestepped by using OGAR on lab-operated computers. If the Client is operated on a lab-operated computer, then near total compatibility can be achieved.

Getting started with OGAR

Individuals interested in using OGAR can view relevant documentation about getting started as well as other details about the project on its GitHub (https://github.com/mboerwinkle/OGAR). Recommendations for server set up, new OGAR releases, community contributions, and other relevant commentary will be updated regularly as the project

continues its development. Interested parties can follow the page to receive notification of any related changes. The authors also welcome correspondence should readers have additional questions about OGAR or require additional support.

Conclusions

Developing virtual alternatives to traditional in-person field research in the arts has the potential to make both basic research and applied assessments of art engagement (e.g., by people working in visitor studies, art education, and museum curation) more affordable, accessible, and safer during public health crises. OGAR may find use with the arts researcher looking for a way to transcend the research-design limitations of physical museum spaces and ever-changing needs of experimental design, with the curator who needs a cost-effective, time-effective way to collect data on curatorial choices for upcoming exhibitions, or with the museum studies class that requires a safe and accessible way for students to engage with gallery spaces without leaving the classroom—all while achieving an acceptable degree of similarity with real-life experiences.

Appendix

Table 4 List of Artworks

Number	Title	Artist	Year	Room	Representational?
1	Lita Curtain Star	Andy Warhol	1968	1	Yes
2	Woman with a Fan	Pablo Picasso	c. 1905	1	Yes
3	Water Lilies	Claude Monet	1906	1	Yes
4	Untitled (Green and Red)	Fiona Rae	1994	1	No
5	Terrano X	Emil Schumacher	1990	2	No
6	Starry Night	Vincent van Gogh	1889	2	Yes
7	Broadway Boogie Woogie	Piet Mondrian	1943	2	No
8	Painting Number 2	Franz Kline	1954	2	No
9	Untitled	Willem de Kooning	1988	2	No
10	Solitary Tree	Caspar David Friedreich	1822	2	Yes
11	Untitled	Per Kirkeby	2012	2	No
12	Eyes in the Heat	Jackson Pollock	1946	1	No
13	Reclining Girl	François Boucher	1752	1	Yes
14	Untitled	Mark Rothko	1954	1	No
15	The Silver Goblet	Jean-Baptiste-Siméon Chardin	1728	1	Yes
16	Hare	Albrecht Dürer	1502	2	Yes

Declarations

Financial interests The authors declare they have no financial interests.

Conflict of interests The authors have no conflicts of interest to declare that are relevant to the content of this article.

Ethics approval Approval was obtained from the Institutional Review Board of the University of North Carolina at Greensboro (Study #21-0311).

Consent to participate Informed consent was obtained from all individual participants included in the study.

References

Agostino, D., Arnaboldi, M., & Lampis, A. (2020). Italian state museums during the COVID-19 crisis: From onsite closure to online openness. *Museum Management and Curatorship, 35*(4), 362–372. https://doi.org/10.1080/09647775.2020.1790029

Augustin, A. (2021). *websockets: An implementation of the WebSocket Protocol (RFC 6455 & 7692)* (9.1) [Python; OS Independent]. Retrieved August 16, 2021, from https://github.com/aaugustin/websockets

Belke, B., Leder, H., Strobach, T., & Carbon, C.-C. (2010). Cognitive fluency: High-level processing dynamics in art appreciation. *Psychology of Aesthetics, Creativity, and the Arts, 4*(4), 214. https://doi.org/10.1037/a0019648

Bennett, T. (2013). *Museums and 'the people'. The birth of the museum: History, theory, politics* (pp. 109–126). Routledge. https://doi.org/10.4324/9781315002668

Bourdeau, L., & Chebat, J. C. (2001). An empirical study of the effects of the design of the display galleries of an art gallery on the movement of visitors. *Museum Management and Curatorship, 19*(1), 63–73. https://doi.org/10.1080/09647770100601901

Brieber, D., Nadal, M., & Leder, H. (2015). In the white cube: Museum context enhances the valuation and memory of art. *Acta Psychologica, 154*, 36–42. https://doi.org/10.1016/j.actpsy.2014.11.004

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

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. Jordan, B. Thomas, & B. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.

Carbon, C.-C. (2017). Art perception in the museum: How we spend time and space in art exhibitions. *I-Perception, 8*(1). https://doi.org/10.1177/2041669517694184

Choi, S. (2013). Relational aesthetics in art museum education: Engendering visitors' narratives through participatory acts for interpretive experience. *Studies in Art Education, 55*(1), 51–63. https://doi.org/10.1080/00393541.2013.11518916

Clarke, J. C., Shortess, G. K., & Richter, M. L. (1984). Stimulus size, viewing distance, and experimental aesthetics. *Visual Arts Research, 10*(2), 1–8.

Cotter, K. N., Fekete, A., & Silvia, P. J. (2021). Why do people visit art museums? Examining visitor motivations and visit outcomes. *Empirical Studies of the Arts.* https://doi.org/10.1177/02762374211011740

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis.* Routledge. https://doi.org/10.4324/9780203807002

Ecma-404. (2017). The JSON Data Interchange Format. Second Edition. http://www.ecma-international.org/publications/standards/Ecma-404.html.

Estrada-Gonzalez, V., East, S., Garbutt, M., & Spehar, B. (2020). Viewing art in different contexts. *Frontiers in Psychology, 11.* https://doi.org/10.3389/fpsyg.2020.00569

Garbutt, M., East, S., Spehar, B., Estrada-Gonzalez, V., Carson-Ewart, B., & Touma, J. (2020). The embodied gaze: Exploring applications for mobile eye tracking in the art museum. *Visitor Studies, 23*(1), 82–100. https://doi.org/10.1080/10645578.2020.1750271

George, A. (2015). *The curator's handbook: Museums, commercial galleries, independent spaces.* Thames & Hudson.

Google LLC (2011). Google arts and culture. https://artsandculture.google.com/

Heidenreich, S. M., & Turano, K. A. (2003). What predicts where one will look when viewing artwork? *Journal of Vision, 3*(9), 688–688. https://doi.org/10.1167/3.9.688

Heidenreich, S. M., & Turano, K. A. (2011). Where does one look when viewing artwork in a museum? *Empirical Studies of the Arts, 29*(1), 51–72. https://doi.org/10.2190/EM.29.1.d

Hipp, R. D. (2021). *SQLite.* Downloaded from https://www.sqlite.org/index.html

Hughes, J. (2021). *reghelper: Helper functions for regression analysis.* R package version 1.0.2. https://cran.r-project.org/web/packages/reghelper/index.html

Ikei, Y., Abe, K., Masuda, Y., Okuya, Y., Amemiya, T., & Hirota, K. (2013, July). Virtual experience system for a digital museum. In: *International Conference on Human Interface and the Management of Information* (pp. 203–209). Springer. https://doi.org/10.1007/978-3-642-39226-9_23

Kennedy, R. S., Stanney, K. M., & Dunlap, W. P. (2000). Duration and exposure to virtual environments: Sickness curves during and across sessions. *Presence Teleoperators and Virtual Environments, 9*(5), 463–472. https://doi.org/10.1162/105474600566952

Keshavarz, B., & Hecht, H. (2012, September). Visually induced motion sickness and presence in videogames: The role of sound. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, No. 1, pp. 1763–1767). Sage. https://doi.org/10.1177/1071181312561354

Khronos Group (2011). WebGL specification. https://www.khronos.org/registry/webgl/specs/1.0.3/

Laramée, F. D. (2002). *Game design perspectives.* Charles River Media.

Leder, H., Gerger, G., Dressler, S. G., & Schabmann, A. (2012). How art is appreciated. *Psychology of Aesthetics, Creativity, and the Arts, 6*, 2–10. https://doi.org/10.1037/a0026396

Leston, J. (1996). Virtual reality: The IT perspective. *ITNOW, 38*(3), 12–13. https://doi.org/10.1093/combul/38.3.12

Lewis, J. R. (2018). The system usability scale: Past, present, and future. *International Journal of Human Computer Interaction, 34*(7), 577–590. https://doi.org/10.1080/10447318.2018.1455307

Lewis, J. R., & Sauro, J. (2009). The factor structure of the System Usability Scale. In M. Kurosu (Ed.), *Human centered design* (pp. 94–103). Springer. https://doi.org/10.1007/978-3-642-02806-9_12

Long, J. S. (1997). *Regression models for categorical and limited dependent variables.* Sage.

Lüdecke, D., Ben-Shachar, M., Patil, I., & Makowski, D. (2020). Extracting, computing, and exploring the parameters of statistical models using R. *Journal of Open Source Software, 5*(53), 2445. https://doi.org/10.21105/joss.02445

Maniaci, M. R., & Rogge, R. D. (2014). Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality, 48*, 61–83. https://doi.org/10.1016/j.jrp.2013.09.008

Mitschke, V., Goller, J., & Leder, H. (2017). Exploring everyday encounters with street art using a multimethod design. *Psychology of Aesthetics, Creativity, and the Arts, 11*(3), 276–283. https://doi.org/10.1037/aca0000131

Moghadam, K., Banigan, C., & Ragan, E. D. (2020). Scene transitions and teleportation in virtual reality and the implications for spatial awareness and sickness. *IEEE Transactions on Visualization and Computer Graphics, 26*(06), 2273–2287. https://doi.org/10.1109/TVCG.2018.2884468

Möller, T., & Trumbore, B. (1997). Fast, minimum storage ray-triangle intersection. *Journal of Graphics Tools, 2*(1), 21–28. https://doi.org/10.1080/10867651.1997.10487468

Mozilla. (2021). *<canvas>: The Graphics Canvas element.* MDN Web Docs. https://developer.mozilla.org/en-US/docs/Web/HTML/Element/canvas.

Nystrom, B. D., & Fairchild, M. D. (1992). Perceived image quality of 16:9 and 4:3 aspect ratio video displays. *Journal of Electronic Imaging, 1*(1), 99–103. https://doi.org/10.1117/12.55178

Patel, H., & Cardinali, R. (1994). Virtual reality technology in business. *Management Decision, 32*(7), 5–12. https://doi.org/10.1108/00251749410068111

Pelowski, M., Forster, M., Tinio, P. P. L., Scholl, M., & Leder, H. (2017). Beyond the lab: An examination of key factors influencing interaction with "real" and museum-based art. *Psychology of Aesthetics, Creativity, and the Arts, 11*(3), 245–264. https://doi.org/10.1037/aca0000141

Pelowski, M., Liu, T., Palacios, V., & Akiba, F. (2014). When a body meets a body: An exploration of the negative impact of social interactions on museum experiences of art. *International Journal of Education & the Arts, 15*(14), n14.

Proctor, N. (2011). The Google Art Project: A new generation of museums on the Web? *Curator: The Museum Journal, 54*(2), 215–221. https://doi.org/10.1111/j.2151-6952.2011.00083.x

R Core Team. (2021). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Reitstätter, L., Brinkmann, H., Santini, T., Specker, E., Dare, Z., Bakondi, F., Miscená, A., Kasneci, E., Leder, H., & Rosenberg, R. (2020). The display makes a difference: A mobile eye tracking study on the perception of art before and after a museum's rearrangement. *Journal of Eye Movement Research, 13*(2). https://doi.org/10.16910/jemr.13.2.6

Revelle, W. (2021). *psych: Procedures for psychological, psychometric, and personality research*. R package version 2.1.6. https://CRAN.R-project.org/package=psych

Rodriguez, R. M., Fekete, A., Silvia, P. J., & Cotter, K. N. (2021). The art of feeling different: Exploring the diversity of emotions experienced during an art museum visit. *Psychology of Aesthetics, Creativity, and the Arts.* Advance online publication. https://doi.org/10.1037/aca0000443

Santini, T., Brinkmann, H., Reitstätter, L., Leder, H., Rosenberg, R., Rosenstiel, W., & Kasneci, E. (2018). The art of pervasive eye tracking: Unconstrained eye tracking in the Austrian Gallery Belvedere. *Proceedings of the 7th Workshop on Pervasive Eye Tracking and Mobile Eye-Based Interaction*, 1–8. https://doi.org/10.1145/3208031.3208032

Smith, J. K. (2014). *The museum effect: How museums, libraries, and cultural institutions educate and civilize society.* Rowman & Littlefield.

Smith, J. K., & Smith, L. F. (2001). Spending time on art. *Empirical Studies of the Arts, 19*, 229–236. https://doi.org/10.2190/5MQM-59JH-X21R-JN5J

Smith, L. F., & Smith, J. K. (2006). The nature and growth of aesthetic fluency. In P. Locher, C. Martindale, & L. Dorfman (Eds.), *New directions in aesthetics, creativity, and the arts* (pp. 47–58). Baywood.

Smith, L. F., Smith, J. K., & Tinio, P. P. L. (2017). Time spent viewing art and reading labels. *Psychology of Aesthetics, Creativity, and the Arts, 11*(1), 77–85. https://doi.org/10.1037/aca0000049

Specker, E., Tinio, P. P. L., & van Elk, M. (2017). Do you see what I see? An investigation of the aesthetic experience in the laboratory and museum. *Psychology of Aesthetics, Creativity, and the Arts, 11*(3), 265–275. https://doi.org/10.1037/aca0000107

Specker, Eva, Stamkou, Eftychia, Pelowski, Matthew, & Leder, Helmut. (2020). Radically revolutionary or pretty flowers? The impact of curatorial narrative of artistic deviance on perceived artist influence. *Psychology of Aesthetics, Creativity, and the Arts*. https://doi.org/10.1037/aca0000320.

Stoffregen, T. A., Faugloire, E., Yoshida, K., Flanagan, M. B., & Merhi, O. (2008). Motion sickness and postural sway in console video games. *Human Factors, 50*(2), 322–331. https://doi.org/10.1518/001872008X250755

Tinio, P. P. L., Smith, J. K., & Smith, L. F. (2015). The walls do speak: Psychological aesthetics and the museum experience. In P. P. L. Tinio & J. K. Smith (Eds.), *The Cambridge handbook of the psychology of aesthetics and the arts* (pp. 195–218). Cambridge University Press. https://doi.org/10.1017/CBO9781139207058.011

Tinio, P. P. L., & Specker, E. (2020). Observation method in empirical aesthetics. In M. Nadal & O. Vartanian (Eds.), *Oxford handbook of empirical aesthetics*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198824350.013.10

Tröndle, M. (2014). Space, movement and attention: Affordances of the museum environment. *International Journal of Arts Management, 17*(1), 4–17.

Voronoï, G. (1908). Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième mémoire. Recherches sur les parallélloèdres primitifs. *Journal fur die Reine und Angewandte Mathematik, 1908*(134), 198–287. https://doi.org/10.1515/crll.1908.134.198

Whitty, M., Cossell, S., Dang, K. S., Guivant, J., & Katupitiya, J. (2010). Autonomous navigation using a real-time 3d point cloud. In: *2010 Australasian Conference on Robotics and Automation* (pp. 1–3).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.