Bird, J. M., Smart, P. A., Harris, D. J., Phillips, L. A., Giannachi, G., & Vine, S. J. (2022). A Magic Leap in tourism: Intended and realized experience of head-mounted augmented reality in a museum context. *Journal of Travel Research.* Advance online publication. https://doi.org/10.1177/00472875221134031

A Magic Leap in Tourism: Intended and Realized Experience of Head-Mounted Augmented Reality in a Museum Context

Abstract

Augmented reality (AR) is an emergent technology in tourism. However, research concerning the AR user experience is relatively scarce and seldom addresses the intentions of designers. Accordingly, we sought to: (a) explore the design intentions underlying a multi-user, purpose-built AR experience; (b) assess the extent to which users' realized experience aligned with the designers' intended experience; and (c) examine the relationships between users' internal states and their associated behavior, in alignment with a Stimulus-Organism-Response framework. In Study 1, designers (n = 5) took part in a focus group and completed a design intentions survey. In Study 2, users (n = 48) tested the AR experience, and a range of subjective (e.g., affective responses) and objective (i.e., visual attention) data were recorded. Findings indicated designer–user disparities primarily at the organism and response levels. Additionally, users' affective responses to the AR experience were strongly associated with visitor engagement.

Keywords: Eye-tracking, mixed methods, mixed reality, remembered utility, stimulus-organism-response, virtual reality.

Resubmitted: September 29, 2022

Introduction

Innovative experiences broaden the visitor appeal and enhance economic performance (Barron & Leask, 2017). This is particularly important in light of the COVID–19 pandemic (Itani & Hollebeek, 2021), given that the travel and tourism industry has suffered losses of almost $4.5 trillion (WTTC, 2021). Technological advancements, such as augmented reality (AR), provide opportunities to enhance the visitor experience by overlaying digital content onto users' immediate surroundings (Loureiro et al., 2020). While discussions concerning the application of AR in tourism span the past two decades (Fritz et al., 2005; Jingen Liang & Elliot, 2021), only recent innovations in hardware and software have enabled AR technology to break out of the laboratory environment into everyday experience.

One sector that has been particularly active in the employment of AR technology is museums (Serravalle et al., 2019). The National Museum of Singapore, for example, provides a *Story of the Forest* smartphone installation depicting the dense tropical rainforests of Southeast Asia and native fauna. The uptake of AR is rapidly accelerating, with the market estimated to be worth $340 billion by the year 2028 (Grand View Research, 2021). Facebook's rebrand to Meta, and the envisioned metaverse, emphasizes the importance of immersive technology that seamlessly integrates physical, digital, and social components. However, it is unlikely that AR's potential will be reached unless there is congruence between designers' intentions and users' actual experiences, and hence there is a need to more fully understand this relationship.

Several research clusters can be identified in the AR tourism literature (He et al., 2018; Jingen Liang & Elliot, 2021). For example, scholars have focused on the challenges and possibilities associated with AR technology (Kounavis et al., 2012), as well as user acceptance (Chung et al., 2015). Conversely, research pertaining to the AR user experience is relatively scarce and seldom addresses the intent of designers who develop such experiences

(Jingen Liang & Elliot, 2021). This is problematic for two reasons. First, little is known about what designers are hoping to achieve by employing AR technology in tourism, with respect to both the type of experience and how it affects users. Second, the extent to which designers' intentions are realized by users of AR technology is presently unknown. A joint approach, which incorporates the perspectives of designers and users, is warranted to comprehensively assess AR experiences in tourism. Identifying the relative points of convergence and divergence between both populations can help inform future theoretical advancements in visitor experience design (Bonfanti et al., 2021).

Preliminary AR user experience research is almost entirely oriented toward the effects of smartphone-based applications that entail a single user engaging with a pre-existing artefact or exhibit (tom Dieck et al., 2018). Importantly, such applications do not encourage interactions with companions (Ponsignon & Derbaix, 2020). The recent development of head-mounted displays, such as the Magic Leap and Microsoft HoloLens, allow multiple users to interact with purpose-built AR experiences in an engaging manner. It is crucial to examine this medium given that social factors are theorized to play a fundamental role in tourism-related experiences (Chen et al., 2020).

This multi-study investigation was conducted within a pragmatic research paradigm (Feilzer, 2010), wherein multiple methods of data collection were employed to address three aims. The aim of Study 1 was to explore the design intentions underlying an AR experience developed for use in a museum context. Study 1 involved the collection of qualitative focus group data and quantitative survey data, in alignment with the suggestion that mixed methods should be prioritized in AR-related tourism research (Jingen Liang & Elliot, 2021). Study 2 involved the collection of quantitative data in relation to user experience testing and this work was guided by the following aims. The primary aim was to assess the extent to which users' realized experience aligned with the designers' intended experience. A secondary aim was to examine the relationships between users' internal states and their associated behavior in relation to the AR experience.

To provide a theoretical foundation for the work, constructs of interest were organized within a Stimulus-Organism-Response (SOR) framework (Suh & Prophet, 2018). SOR models have been applied to the study of virtual reality (VR) technology and thus the present study provides a logical extension to this line of inquiry (Kim et al., 2020). The SOR model employed in the present investigation provided equal coverage of cognitive (e.g., presence) and affective (e.g., remembered pleasure) constructs. This is noteworthy given that many conceptual frameworks in the tourism domain do not dedicate sufficient attention toward affective phenomena (McCabe et al., 2016; Tucker & Shelton, 2018).

From a practical perspective, the involvement of a creative team afforded an opportunity to identify and establish design intentions, against which user experience testing could subsequently be compared (Smit et al., 2021). We collected user data pre-, during, and post-experience to examine how visitor responses unfolded over time (Stienmetz et al., 2021). Eye-tracking data was captured to explore users' attentional processes. Such data is less susceptible to bias when compared to subjective data (Scott et al., 2019) and this is one of the first attempts in tourism to incorporate eye-tracking in AR. Finally, the application under investigation is a purpose-built, multi-user experience that requires an AR head-mounted display.

The following sections contain a review of literature and an overview of the developed hypotheses. Study 1 presents an exploration of the design intentions underlying the AR experience, while Study 2 entails user experience testing. Thereafter, a general discussion is offered, which includes theoretical and managerial implications, as well as limitations and directions for future research.

Literature Review

Intended and Realized Experience

Providing memorable experiences is fundamental for the longevity of tourism providers (Barron & Leask, 2017). Accordingly, there is growing interest toward the role of design in tourism (Smit et al., 2021). There is substantial variation in the conceptualization of tourism experience design (Tussyadiah, 2014). Herein, we conceptualize experience design as the practice of designing a product or service with an emphasis on the quality of user experiences. Hence, this process is predicated on the creation and staging of prerequisites that enable consumers to have desired experiences (Smit et al., 2021).

A distinction can be made between intended and realized experiences (Voss et al., 2008). *Intended* experiences are planned by tourism providers and offered to visitors. Alternatively, *realized* experiences refer to the actual lived experiences of visitors. Concentrating on the perspectives of visitors alone does not provide sufficient intellectual insight to inform experience design. Consequently, researchers have advocated a joint approach that considers the views of designers and visitors in combination to better inform tourism-related experience design (Dube et al., 2015).

There are many advantages associated with the inclusion of designers in the research process. For example, designers can identify constructs that were important to them during development and hence their views can serve as a foundation upon which to compare users' experiences (Ponsignon et al., 2017). Additionally, the findings from user testing can be communicated back to designers so that they can refine the experience in accordance with an iterative design process (Tussyadiah, 2014). This would afford an opportunity to minimize any designer–user disparities, thereby enhancing the visitor experience and maximizing the site's economic performance.

Few attempts have been made to assess the congruence between intended and realized experience and this is a focus of the present investigation. A notable exception concerns a qualitative study by Ponsignon et al. (2017). The researchers conducted a series of interviews with visitors and staff at a cultural center and identified four areas that can be managed by the service provider (i.e., individual touchpoints, customer journey, social environment, and physical environment). Nonetheless, there is significant scope to employ quantitative approaches to the study of intended and realized experience. This was emphasized by Ponsignon et al. (2017), who recommended that visitor behavior data be captured in real-time. Head-mounted AR devices collect real-time behavioral data through a series of integrated sensors. Hence, there appears to be a strong rationale for the study of head-mounted AR in tourism.

Stimulus-Organism-Response Models

SOR models assert that external or environmental cues prompt cognitive and affective states, which subsequently drives behavioral responses (Jacoby, 2002; Mehrabian & Russell, 1974). Tourism researchers have employed SOR models to help explain consumer behavior in relation to destinations shown in 360-degree images (Yeh et al., 2017) and VR (Kim et al., 2020). The present investigation extends this line of scientific inquiry to incorporate AR technology. A SOR model is used as a guiding framework to explore the relationships between the AR experience (stimulus), users' internal states (organism), and their associated behavior (response), and is also used to assess the extent to which users' realized experience aligns with designers' intended experience. Fundamental constructs related to each phase of the SOR model are depicted in Figure 1 and described herein.

*****Insert Figure 1 about here*****

Stimulus. The *stimulus* refers to the trigger that facilitates users' cognitive and affective reactions. Suh and Prophet (2018) made the distinction between technological stimuli and content stimuli. The technological stimulus under investigation is head-mounted AR, which overlays digital assets onto users' real environments. An important feature of AR in tourism is the ability for users to access engaging digital content while concurrently preserving interactions with other users, thereby adding a social element to the stimulus. This is beneficial because social interaction is a vital component of experience that has been associated with a range of desirable outcomes (e.g., positive affective responses; Chen et al., 2020). In contrast, VR, despite being fully immersive, can be considered to prompt rather solitary experiences (Ingram et al., 2019). It is perhaps for this reason, that AR has been described as one of the most promising technologies in tourism (Loureiro et al., 2020).

It is also important to consider the type of content or experience that is delivered via AR technology, as different applications will prompt diverse affective and cognitive states (Suh & Prophet, 2018). Experience typologies are useful for tourism managers, as they afford the distinction between different market offerings. A plethora of visitor experience typologies have been proposed (Packer & Ballantyne, 2016), with Pine and Gilmore's (1998) Experience Economy serving as a predominant framework in tourism (Jung et al., 2016).

Pine and Gilmore (1998) emphasized the importance of staged experiences and proposed four realms according to two dimensions: involvement (ranging from passive to active participation) and desire (ranging from absorption to immersion). *Active participation* permits the visitor to directly affect the event or performance (e.g., attending an interactive cooking class), however, such interaction is not possible during *passive participation* (e.g., watching a film at a cinema; Pine & Gilmore, 1998). *Absorption* occurs when an experience occupies an individual's attention whereas *immersion* refers to instances where an individual feels part of the experience itself (Pine & Gilmore, 1998). Subsequently, the researchers proposed four realms of experience (i.e., entertainment, education, aesthetics, and escapism) and suggested that the richest experiences comprise elements from all four realms. A strength of the experience economy framework is its flexibility in identifying a range of experience types. Nonetheless, there is also a need to consider distinct psychological constructs at the organism level when studying tourism experiences (Scott & Le, 2017).

Organism: Cognitive. Organism refers to users' internal evaluations of the stimulus. It is possible to identify a range of cognitive reactions to immersive technology use (Suh & Prophet, 2018). In the context of AR technology, *presence* is a sense of feeling surrounded by a realistic physical/virtual environment (Georgiou & Kyza, 2017). Presence in AR has been associated with a range of positive outcomes such as greater intentions to use smartphone-based shopping applications (Smink et al., 2020), but has rarely been the subject of empirical investigation in tourism (Jung et al., 2016). A notable exception concerns a study by He et al. (2018), who reported that greater perceptions of presence were associated with increased willingness to pay for an art museum experience. However, the researchers' intervention was pre-recorded, and participants were instructed to imagine that they were on-site using an AR device. Therefore, there is ample opportunity to examine presence in AR with greater consideration toward ecological validity.

Another cognitive process of interest concerns *visual attention* (Scott et al., 2019), which allows individuals to selectively prioritize or suppress information in their environments. Understanding visitors' visual attention is important to site managers, as this information can be used to help redesign staged experiences in an engaging manner (Le et al., 2020). Our eyes are constantly in motion of two main types. *Fixations* refer to instances in which the eyes remain still and can last anywhere from tens of milliseconds to several seconds. Alternatively, rapid movements of the eyes from one fixation to another are termed *saccades* and can take 30–80 ms to complete (Holmqvist & Andersson, 2017). Fixations reveal the stimuli that visitors are dwelling on whereas saccades are indicative of a shift in focus.

Eye-tracking technology has become a powerful tool to help researchers examine the perception of visual stimuli in tourism (Rainoldi & Jooss, 2020). For example, screen- and mobile-based eye-tracking devices have been used to examine tourists' attentional processes in relation to marketing materials (Scott et al., 2019). However, there is a dearth of eye-tracking research in relation to immersive virtual environments, despite researchers emphasizing the usefulness of such an approach (Rainoldi & Jooss, 2020). This paucity of research is surprising given that many head-mounted displays used to access immersive stimuli have the capacity to record users' gaze behavior (Harris, Bird, et al., 2020).

Organism: Affective. Conceptual frameworks in tourism often support the notion that visitors are rational decision makers (McCabe et al., 2016). Visitors are typically theorized to engage complex cognitive processes when collecting, rationally evaluating, and acting upon information that serves to promote their greatest satisfaction (Pearce & Packer, 2013; Walls et al., 2011). However, visitors do not always make fully rational decisions and their motives are often driven by the powerful influence of affective phenomena (Walls et al., 2011; Wattanacharoensil & La-ornual, 2019). Affective responses are increasingly being cited as key determinants of memorable experiences and hence their measurement is important (Chen et al., 2020; Godovykh & Tasci, 2020a).

Despite the appeal of measuring affective responses in tourism, the terms affect, emotion, and mood are often used interchangeably reducing conceptual clarity (Skavronskaya et al., 2017). Accordingly, there is a need to define constructs of interest if this line of scientific inquiry is to flourish (Skavronskaya et al., 2017). *Affect* can be defined as "a neurophysiological state consciously accessible as a simple primitive nonreflective feeling most evident in mood and emotion but always available to consciousness" (Russell & Feldman Barrett, 2009, p. 104).

Herein, affect is conceptualized as a dimensional domain, containing two orthogonal and bipolar dimensions, affective valence (ranging from *pleasure* to *displeasure*) and arousal (ranging from *sleepiness* to *high arousal*). Researchers have advocated the measurement of affective valence, albeit adopting the behavioral economics term *utility* (Kahneman et al., 1997), as an appropriate successor to satisfaction (Godovykh & Tasci, 2020b). This is because affective valence is a bipolar dimension and has greater coverage of outcomes when compared to satisfaction, which is typically assessed with unipolar measures.

In addition to measuring affective responses during a specific encounter, Godovykh and Tasci (2020b) recommended the measurement of affective responses in relation to two additional timepoints. *Remembered pleasure*/*utility* concerns how pleasant or unpleasant an experience is later remembered. Moreover, *forecasted pleasure/utility* concerns how pleasant or unpleasant future experiences are predicted to be. Researchers have theorized that both remembered and forecasted pleasure can help predict whether behavior will be repeated (Karl et al., 2021; Zenko et al., 2016). Hence, such constructs could yield considerable value when evaluating the extent to which AR experiences can help retain visitors.

Few investigators have sought to assess affective responses to AR experiences in the tourism domain. Kourouthanassis et al. (2015) conducted a field study with visitors using a smartphone-based travel guide. Participants were required to use the application for the duration of their visit and the researchers reported that affective valence and arousal were both statistically significant predictors of usage behavior. However, little is known about the effects of AR experiences on remembered/forecasted pleasure. Hence, there is a distinct need to consider this element of the user experience and how it relates to visitor behavior.

Organism: Individual differences. There is evidence to suggest that individual differences can influence responses to immersive technology use (Suh & Prophet, 2018). For example, Park and Stangl (2020) reported that high-sensation seekers reported the most positive AR experiences when compared to their lower sensation-seeking counterparts. Another factor that might influence responses to immersive technology use concerns the social environment (Bolton et al., 2018; Ponsignon & Derbaix, 2020). Preliminary findings from computer science indicates that shared AR experiences can prompt greater user interest when compared to individual applications (Park et al., 2020). However, future research is required to examine the extent to which this applies in tourism.

Responses. Responses refer to the outcomes of immersive technology use (e.g., learning effectiveness; Suh & Prophet, 2018). Perhaps the most important response to museums is visitor engagement (Barron & Leask, 2017). This can be conceptualized as involvement with, and commitment to, a consumption experience (Taheri et al., 2014). Researchers have shown that AR applications can prompt visitor engagement in relation to science festivals (tom Dieck et al., 2018), but further research is required to assess this in a museum context.

It is important to consider positive and negative responses to immersive technology use (Suh & Prophet, 2018). A negative response is cognitive overload, given that tourism experiences typically entail an element of visitor learning. In accordance with Cognitive Load Theory (Sweller, 1999), cognitive workload can be considered as the quantity of informational units that must be held in working memory during a task. It is plausible that there is an optimal level of cognitive workload for tourism-related AR experiences, and it is the responsibility of designers to optimize this. Previous research is heavily weighted toward positive user responses (Kim et al., 2020; tom Dieck et al., 2018). Hence, examining negative responses to AR experiences, such as cognitive workload, represents a more harmonious approach when compared to the extant literature (Jingen Liang & Elliot, 2021).

Hypothesis Development

Intended and Realized Experience

AR design and user experience studies have been described as promising lines of inquiry (Jingen Liang & Elliot, 2021). However, no research to date has addressed the extent to which designers' intentions are realized by users of AR technology in tourism. Study 1 comprised an exploration of the design intentions underlying an AR experience developed for a museum context. A mixed methods approach was adopted and included the collection of qualitative focus group data and quantitative survey data. Study 1 was exploratory (no a priori hypotheses) and generated predictions that were assessed in Study 2, which entailed the collection of quantitative data in relation to user experience testing. Accordingly, specific research hypotheses pertaining to intended and realized experience are presented in Study 1, following data collection with designers.

User Responses

An aim of Study 2 was to examine the relationships between users' internal states and their associated behavior in relation to the AR experience. Constructs of interest were organized in accordance with a SOR framework (see Figure 1). Cognitive and affective reactions to immersive stimuli are theorized to prompt user responses (Suh & Prophet, 2018). Hence, we hypothesized positive relationships between presence–visitor engagement and between visual attention–visitor engagement (H_1). Likewise, we predicted positive relationships between remembered pleasure–visitor engagement, and between forecasted pleasure–visitor engagement (H_2). Kim et al. (2020) sought to examine tourists' intentions to visit destinations previously experienced in VR. The researchers employed a SOR framework and reported that cognitive factors had a stronger influence on positive outcomes when compared to affective responses. Accordingly, we predicted stronger relationships between cognitive variables–visitor engagement when compared to the relationships between affective variables–visitor engagement (H_3).

The majority of AR research in tourism concerns smartphone-based applications that are designed for a solitary user (tom Dieck et al., 2018; Trunfio & Campana, 2020).

Consequently, little is known about the effects of engaging with purpose-built, head-mounted AR experiences among groups of visitors. Initial findings indicate more positive responses to shared experiences when compared to individual experiences (Park et al., 2020). Hence, we hypothesized greater visitor engagement for those who take part in a shared experience when compared to individual users (H_4). At the organism level, we predicted that shared experiences would prompt more positive remembered and forecasted pleasure when compared to individual users (H_5). Conversely, we hypothesized that shared experiences would be associated with lower presence and visual attention toward digital assets, when compared to individual users (H_6). This is because cognition is likely to gravitate toward other users during shared experiences.

Study 1

Methods

Pragmatism. The present investigation was conducted in alignment with a pragmatic research paradigm (Feilzer, 2010). A detailed description of this approach can be found in Supplementary Material 1.

Participants. The study was approved by the University of Exeter Research Ethics Committee. A purposive sample of five adult designers was recruited through email correspondence with a production studio in the UK (M_{age} = 44.8 years, SD_{age} = 3.2 years; one woman, three men, and one who preferred not to say; four British and one who preferred not to say; $M_{\text{experience}}$ = 20.6 years; $SD_{\text{experience}}$ = 2.7 years). Inclusion criteria stipulated that volunteers were involved in the creation of the AR application. Each studio department (e.g., design, quality assurance) were represented. It was anticipated that the sample size was small enough for each participant to contribute, yet sufficiently large to share diverse opinions across the whole group (Freeman, 2006).

Procedure. Data collection took place in May, 2021. A convergent mixed methods design was employed, which entailed a concurrent collection and analysis of qualitative and quantitative data (Creswell, 2022). This research design was appropriate given that it allowed for the perspectives of designers to emerge via qualitative methods, but also allowed us to quantify designers' intended experience, enabling subsequent analyses against users' realized experience in Study 2. Participants provided informed consent and completed a demographic questionnaire. A review of literature was conducted to facilitate the development of focus group materials (He et al., 2018; Suh & Prophet, 2018), which were subsequently refined among the research team. The purpose of the focus group was to explore the design intentions underlying the AR experience. The focus group was conducted via web-based videoconferencing software (Zoom; San Jose, CA, USA) and all participants joined using video and audio. It is important to create a comfortable environment when conducting focus groups (Krueger & Casey, 2014). Zoom was deemed appropriate given that (a) participants could join the discussion from a familiar setting, (b) participants could easily see and hear other participants, and (c) it could limit in-person interaction during the COVID–19 pandemic.

Three members of the research team were present during the focus group. The lead researcher served to moderate the discussion while the remaining two researchers provided additional assistance (e.g., monitoring the chat function on Zoom; Krueger & Casey, 2014). The focus group commenced with a brief introduction from the moderator, who reiterated the purpose of the study and informed participants that the discussions would serve to guide the next phase of the research.

The opening question concerned each participant's role in the development of the AR experience. This opening question was intended to be easy to answer to encourage participants to divulge information and feel comfortable in the group setting (Krueger &

Casey, 2014). The second question related to the origins of the project and participants were encouraged to contribute until no more views were offered. The main topics of discussion were design challenges, target user demographics, as well as user responses during and post-experience. The questions were open-ended and intended to encourage universal participation within the group. The moderator used the screen share function on Zoom to display a series of documents (e.g., internal storyboards) to facilitate a critical discussion of the AR experience on a scene-by-scene basis.

The moderator employed follow-up questions and provided opportunities to clarify responses to explore the subject at a deeper level. The focus group lasted 61 min, was digitally recorded, transcribed verbatim, and yielded 17 pages of single-spaced text. Each participant was required to complete a design intentions survey via web-based software (Qualtrics; Provo, UT, USA) upon the cessation of the focus group. The survey was completed by each member of the design team individually. It was anticipated that this would eliminate any bias occurring through group dynamics, which is a concern when sampling pre-existing groups of individuals who work closely together (Freeman, 2006). The data derived from the survey served as a foundation upon which users' realized experiences could be compared against in Study 2.

Measures. The design intentions survey consisted of multiple inventories designed to capture user responses to AR technology. In all cases, the stem of each item was minimally adjusted to capture design intentions. For example, "The setting of the AR experience was very attractive" was adjusted to "We intended the setting of the AR experience to be very attractive" (tom Dieck et al., 2018).

An inventory developed by tom Dieck et al. (2018) was adapted and employed to assess the stimulus content in relation to the four realms of experience (i.e., entertainment, education, aesthetics, escapism) advocated by Pine and Gilmore (1998), as well as user engagement. This inventory included 17 items (e.g., “We intended for users to learn something new during the AR experience”) attached to a 5-point bipolar scale (1 = *Strongly Disagree,* 5 = *Strongly Agree*).

Presence was measured using 4-items (e.g., “We intended for users to be so involved, that they feel their actions could affect the activity”) adapted from the Augmented Reality Immersion questionnaire (Georgiou & Kyza, 2017). Items were attached to a 7-point bipolar scale (1 = *Totally Disagree,* 7 = *Totally Agree*).

On the basis that we conceptualize affect as a dimensional domain, intended affective responses were firstly assessed using the Affect Grid (Russell et al., 1989). This is a 9 by 9 grid, with the horizontal dimension representing affective valence (from *unpleasantness* to *pleasantness*) and the vertical dimension representing arousal (from *sleepiness* to *high arousal*). Anchors are placed at the extremes of the two orthogonal dimensions (e.g., “pleasant feelings”), as well as the four corners (e.g., “excitement” [pleasant, high-arousal]) to facilitate understanding (Ekkekakis, 2013). The moderator shared the Affect Grid during the focus group and participants were required to collectively select one of the 81 squares that corresponded with their intended experience on a scene-by-scene basis.

To minimize common method variance, remembered pleasure was measured using a scale with a different format to the Affect Grid (Russell et al., 1989). A Visual Analogue Scale was employed in relation to the question “Overall, how did you intend to make users feel during the AR experience?” The scale ranged from -100 (*very unpleasant*) to 100 (*very pleasant*) in intervals of 1. The slider was initially positioned at the origin (0). The descriptors and slider were visible to participants but the numbers were not (Zenko et al., 2016).

Forecasted pleasure was measured using the Empirical Valence Scale (Lishner et al., 2008). Participants were required to respond to the question “If users repeated the AR experience, how do you think they would feel?” Fifteen empirically spaced verbal descriptors were depicted underneath the scale, ranging from -100 (*most unpleasant imaginable*) to 100 (*most pleasant imaginable*). The values were hidden from participants who were instructed to select one descriptor only.

Cognitive workload was measured using an adapted Simulation Task Load Index (SIM-TLX; Harris, Wilson, et al., 2020). This inventory comprised nine items (e.g., “How mentally fatiguing did you intend the task to be?”) attached to a 21-point bipolar scale (0 = *Very Low*, 20 = *Very High*).

Data analysis. Separate analyses were conducted with respect to the qualitative and quantitative data, in alignment with a convergent mixed methods design (Creswell & Plano Clark, 2017). The transcription data were organized using NVivo (QSR; Melbourne, Australia) and analyzed by means of theoretical thematic analysis (Braun & Clarke, 2006). Initially, the lead author engaged in a process of familiarization to gain a sense of the overall context of the data and the wording employed by participants. The transcript was read multiple times and initial ideas were recorded. Thereafter, initial codes were generated theoretically in relation to Suh and Prophet’s (2018) SOR model of immersive technology use. Codes were then collated into larger themes and a reviewal process enabled the development of a thematic map (Terry et al., 2017). Finally, the themes were defined and named, which helped to tell the overall story of the analysis (Maguire & Delahunt, 2017). Responses to the design intention survey were collated and descriptive statistics were calculated. Integration was achieved by merging the results from the qualitative and quantitative data, which enabled a more complete understanding of the designers’ intended experience when compared to that provided by either the qualitative or quantitative results alone (Creswell & Plano Clark, 2017).

Results and Discussion

Stimulus. Participants described several properties pertaining to the AR experience. The decision to employ AR technology from Magic Leap appeared to be predicated on a desire to develop for "one of the boundary pushers in terms of entertainment" (Quality Assurance Lead). The AR experience depicted "ghost dinosaurs … and you've got to try and spot them, and you've got to try and help save them and return them to their dimension" (Creative Director). Hand gestures, which were tracked by the AR technology, appeared to be of central importance: "You would hold out your hand, a stream would go out to a ghost dinosaur, it would encapsulate it and then it would float up and through a portal and back to ghost land" (Lead Designer; See Figure 2).

*****Insert Figure 2 about here*****

The AR experience was developed by a consortium which included two museums. Accordingly, the authenticity of the simulated content was vital: "Even though they were ghosts, they were all scientifically accurate ghost dinosaurs and then we talked about what their species were and what time period they were from" (Lead Designer). This "known facts" (Bec et al., 2019, p. 118) approach draws upon validated information to present an accurate account of history, which is fundamental for visitor education (Mura et al., 2017). The extent to which the required hand gestures aligned with the core values of the corporate partner were also considered:

> We weren't *Ghost Busters;* we weren't destroying them [dinosaurs]. It was really important to the museum … their ethos of saving species… we didn't want it to be a zapping game… we want to feel like we're saving them [dinosaurs], we're not attacking them. (Creative Director)

Participants revealed that they wanted the physical set to represent "a slightly unnerving retro café that would look like it had been attacked by something" (Creative

Director). Unfortunately, “the full café did not get designed” as this stripped back physical set was scheduled to be showcased overseas at a large film festival. The qualitative insights were largely corroborated by the findings from the design intentions survey. Using Pine and Gilmore’s (1998) experience typology as a guiding framework, the AR experience was intended to provide entertainment (*Mdn* = 4.66) and education (*Mdn* = 3.50) to a greater extent than aesthetics (*Mdn* = 3.00) and escapism (*Mdn* = 2.25). Accordingly, we predict statistical equivalence between designers’ intentions and users’ realized experience for each realm of experience (H_7).

Organism: Cognitive. The Lead Designer explained that the AR experience commenced with a small dinosaur that “just appears on the surface in front of you.” As the experience progresses, multiple dinosaurs appear and “they are moving around and interacting with the surfaces, they’re standing on the worktop, they’re standing on the boxes… every bit of physical set has a digital twin.” Presence (i.e., a sense of feeling surrounded by a realistic physical/virtual environment; Georgiou & Kyza, 2017) was clearly of significance to the design team “…if people didn’t get the feeling that they [dinosaurs] were actually interacting with the physical assets then that would not be completely hitting our target” (Lead Designer). This qualitative finding was substantiated by the design survey, which indicated that the experience was intended to prompt high perceptions of presence (*Mdn* = 6.00) and we hypothesize statistically equivalent scores among users (H_8).

The analysis revealed that the peak of the experience came toward the end when a Tyrannosaurus Rex “peeks through a hole in the wall” (Art Lead). There was a high degree of expectancy that this part of the experience should comprehensively capture the visual attention of users, with a Programmer stating that “we would have failed if people were not aware or looking at the T-Rex”. Given these qualitative insights, we predict statistical equivalence between designers' intentions and users' realized experience in relation to visual attention toward the digital Tyrannosaurus Rex (H_9).

Organism: Affective. Designers revealed that the process of equipping an AR head-mounted display in a social environment could raise users' arousal: "there's the potential that one headset doesn't do what it's supposed to do, so there's a feeling of 'Oh, I'm holding the group back', there is a peer pressure almost implied" (Quality Assurance Lead). Participants explained that the early phase of the experience was designed to elicit affective responses in the pleasant, high-arousal quadrant of the Affect Grid (Russell et al., 1989): "It would be quite exciting wouldn't it? There's nothing to be scared of, it's quite a cute dinosaur for the first one, isn't it? And it should feel quite exciting and pleasant" (Creative Director). The Lead Designer added:

> The only time that we would go toward unpleasant feelings is probably emotions like fear and shock, which does feature in this [experience], particularly the T-Rex… there's verbal 'OMGs' and stepping back and it's a punchy moment when it sticks it's head through the wall and the audio that goes with it is really strong.

Participants explained that although not a design intention, the cessation of the experience often results in pleasant affective responses: "it is not intentionally an exciting or enjoyable moment of handing something [AR head-mounted display] back to somebody, but in practice that is what it tends to end up being" (Lead Designer). The affective journey associated with the AR experience is depicted in Figure 3. We predict an increase in users' affective valence and a decrease in users' arousal from pre- to post-experience, in alignment with designers' intentions (H_{10}).

*****Insert Figure 3 about here*****

High scores were observed for the remembered pleasure item of the design intentions survey (*Mdn* = 86.00). However, the forecasted pleasure item yielded lower scores (*Mdn* = 70.00). Hence, it is plausible that the AR experience was designed for single, as opposed to repeated, consumption. We predict statistical equivalence between designers' intentions and users' realized experience for remembered pleasure and forecasted pleasure (H_{11}).

Organism: Individual differences. Participants described the difficulties associated with developing an AR experience for a broad demographic. Devising the gameplay in accordance with the theorized demographics' familiarity with AR appeared to help in this regard:

> We had to make sure that it was extremely accessible for people who had no prior use of video games, the fact that people were wearing a Magic Leap for the first time and could easily spend the first 30 seconds just going 'Woah, what am I looking at?' (Lead Designer)

The AR experience catered for small groups of simultaneous users and there was evidence to suggest that shared experiences would influence visual attention:

> As you go through the dinosaur experience, people work out really quickly that they can steal dinosaur evidence from other people and so you start to see, and I've experienced it myself, you look over and see what other people are doing and sort of going 'Oh right, I'm going to grab that', the competitive nature takes over. (Quality Assurance Lead)

Responses. The design survey revealed that a high level of visitor engagement (e.g., interacting with other dinosaur related materials) was intended following the cessation of the AR experience (*Mdn* = 4.00) and we hypothesize statistically equivalent scores to be reported by users (H_{12}). The designers also explained that they tried to minimize negative responses for users: "Any cognitive load for them [users] beyond really simple interfaces and a really clear objective and they were going to spend the whole 5 min just staring around and going ‘Look at that!’” (Lead Designer). Examination of the SIM-TLX scores confirmed that designers intended the AR experience to prompt low levels of cognitive workload (*Mdn* = 3.00). Temporal demands yielded the highest designer scores (*Mdn* = 8.00), and this was perhaps due to the relatively short duration of the experience:

> Not including the on-boarding and off-boarding on either side, we aimed for 5 mins, so there was some light narrative touch but also when it came to the gameplay, we needed to be able to teach gameplay that people could get really quickly. (Lead Designer)

Bearing these findings into consideration, we hypothesize statistically equivalent cognitive workload scores from users, with temporal demands prompting the highest scores (H_{13}).

Study 2

Methods

Participants. Ethical approval was granted by the University of Exeter Research Ethics Committee. Sample size was determined by a resource constraints approach (i.e., access to the physical set being restricted to a one week period; Lakens, 2022). A purposive sample of 48 adults was recruited (M_{age} = 28.69 years, SD_{age} = 10.64 years; 27 women, 21 men). Recruitment was conducted through word-of-mouth and facilitated by means of social media posts. Inclusion criteria stipulated that participants were 18 years of age or older without visual or auditory impairment that was not corrected for (e.g., with contact lenses). Volunteers were required to provide evidence of a negative COVID–19 lateral flow test prior to participation. Furthermore, volunteers were informed that their participation would enable entry into a raffle, which comprised five £50 gift vouchers. A sensitivity analysis was conducted in R Studio (2022.07.1) to determine the smallest effect size of interest (SESOI) in relation to a one-sample equivalence test (Lakens, 2022). Given $N = 48$, $SD = 1$, and $\alpha = .05$, 80% power was achieved with equivalence bounds ± .42 expressed in raw scores.

Apparatus. The physical set associated with the AR experience was assembled prior to data collection (see Figure 4). This included a semi-circular desk positioned perpendicular to a large grey wall. Several boxes were positioned on the desk that served as props for the digital content to interact with. The AR experience could accommodate up to six simultaneous users (depicted by the colored squares; see Figure 4). However, the research team restricted the number of concurrent users to a maximum of three, to maintain adequate social distancing during the COVID–19 pandemic.

*****Insert Figure 4 about here*****

AR head-mounted displays (Magic Leap 1; Plantation, FL, USA) were used to deliver the digital experience and to record participants' gaze behavior. The AR device consisted of a lightweight headset tethered to a small battery pack. Additionally, handheld controllers were used to navigate through menus. Cleanbox technology (CX1; Carlsbad, CA, USA) and disinfectant wipes were employed to ensure that each AR head-mounted display and controller were thoroughly cleaned between uses.

Procedure. Data collection took place in May, 2021. A cross-sectional study design was employed. Participants visited the site on one occasion to take part in the AR experience. Following COVID–19 checks, volunteers read an information sheet and provided informed consent. Thereafter, they completed a demographic questionnaire. Members of the research team demonstrated how to correctly fit and adjust the AR head-mounted display. Handheld controllers were also provided, and their functions described. Each participant was asked to stand in position around the semi-circular desk. Subsequently, volunteers completed a visual calibration of the AR head-mounted display. This process required participants to fixate on a total of 14 targets presented at a range of locations/depths and served to enhance the validity of the eye-tracking data. Upon successful calibration, volunteers took part in the AR experience which lasted approximately 5 mins. Participants were required to “release” and “collect” evidence from the digital dinosaurs depicted via the AR head-mounted display. Following completion of the experience, participants were instructed to complete a post-experience survey.

Measures. Core affect was assessed pre- and post-experience using the Affect Grid (Russell et al., 1989). All measures contained in the post-experience survey echoed those of the design intention survey, but without adjustment to the stem of each item. The stimulus content was measured using items developed by tom Dieck et al. (2018) in relation to Pine and Gilmore’s (1998) four realms of experience. Presence was assessed by the Augmented Reality Immersion questionnaire (Georgiou & Kyza, 2017). Remembered pleasure and forecasted pleasure were measured using visual analogue scales (Lishner et al., 2008; Zenko et al., 2016). User engagement was assessed using items derived from tom Dieck et al. (2018). Furthermore, cognitive workload was measured using the SIM-TLX (Harris, Wilson, et al., 2020). Additional details (e.g., anchors) are presented in Study 1 and all items are contained in Supplementary Material 2.

Data analysis. Supplementary Material 3 describes the data screening associated with the objective eye-tracking data. The survey data were screened for univariate outliers in R Studio (2022.07.1) using standardized z-scores ($z > \pm 3.29$; Tabachnick & Fidell, 2019). Tests revealed six outliers and in all instances, the score was adjusted by assigning the outlying cases a raw score that was one unit smaller or larger than the next most extreme score in the distribution until $z < \pm 3.29$ (Tabachnick & Fidell, 2019). The distributional properties of the data were examined visually by means of normal Q–Q plots and histograms (Coolican, 2018).

Tests of the distributional properties of the data revealed violations of normality in 15 of the 23 cells of the analysis (three at $p < .05$, four at $p < .01$, and eight at $p < .001$). Scholars have raised concerns about the transformation of subjective data derived from Likert scales (Nevill & Lane, 2007). Hence, these data were not transformed. Subsequently, non-parametric analyses were employed. Such analyses were deemed appropriate given that skewness values frequently exceeded twice the standard error of the dependent variables (Coolican, 2018; see Supplementary Material 4). User responses were assessed by means of Spearman's rho correlations and Wilcoxon rank sum tests. Holm-Bonferroni corrections were applied to help control family-wise error and significance was accepted at $p < .05$.

One-sample Wilcoxon signed rank tests were used to examine statistical equivalence between designers' intentions and users' realized experience. This procedure involved conducting two one-sided tests (TOSTs) to determine whether the location shift was sufficiently close to zero to reject the presence of a meaningful difference. The SESOI was used to set symmetrical equivalence bounds around designers' intended experience in raw scores (e.g., ± .42 on a 5-point scale, ± .59 on a 7-point scale). Statistical equivalence was established when the larger of the two *p* values was smaller than alpha (.05; Lakens et al., 2018). All analyses were conducted in R Studio (2022.07.1) and the associated markdown files are available online (https://osf.io/BT3UV/).

Results and Discussion

All research hypotheses associated with the present investigation are reiterated in Supplementary Material 5.

User responses. Spearman's rho correlations were used to examine the relationships between visitor engagement and presence, visual attention, remembered pleasure, and forecasted pleasure. A moderate positive relationship was observed between presence and visitor engagement (r_s =.38, n = 48, p < .01; see Figure 5a), providing partial support for H_1.

However, visual attention toward digital assets was not associated with visitor engagement ($p > .05$; see Figure 5b). This was a rather unexpected finding that opposes the predictions of SOR models (e.g., Suh & Prophet, 2018). Large positive relationships were observed between remembered pleasure–visitor engagement ($r_s = .52$, $n = 48$, $p < .001$; see Figure 5c) and between forecasted pleasure–visitor engagement ($r_s = .49$, $n = 48$, $p < .001$; see Figure 5d), leading to the acceptance of H_2.

H_3 was not accepted given that the relationships between affective variables–visitor engagement were stronger than those between cognitive variables–visitor engagement. These findings oppose recent VR-related research (Kim et al., 2020). Nonetheless, the results contribute toward a growing corpus of work that emphasizes the importance of affective phenomena in tourism (Godovykh & Tasci, 2020a).

Wilcoxon rank sum tests were employed to examine the effects of experience type (i.e., individual vs. shared) on presence, visual attention, remembered pleasure, forecasted pleasure, and visitor engagement. The analyses indicated that the differences were negligible and statistically non-significant (ps > .05; see Supplementary Material 6). Researchers have frequently suggested that social interaction is integral for desirable outcomes in tourism (Chen et al., 2020; Wei et al., 2019). Hence, the present findings were somewhat unexpected and prohibited the acceptance of $H_{4\text{-}6}$. It is possible that the AR experience was not of sufficient length (i.e., 5 min) to induce the hypothesized differences between individual and shared AR experiences.

*****Insert Figure 5 about here*****

Intended and realized experience.

Stimulus. The TOST procedure (SESOI = .42) indicated statistical equivalence between designers' intended experience and users' realized experience for entertainment ($p = .004$) and education ($p = .001$). However, users' aesthetics and escapism scores were not equivalent to designers' intentions (*p*s > .05; see Figure 6), which prevented the full acceptance of H_7. It is noteworthy that such discrepancies are not inherently negative, as users' scores for aesthetics and escapism surpassed those of the designers (see Figure 6). This is particularly encouraging given that rich experiences are theorized to comprise elements from all four realms (Pine & Gilmore, 1998).

*****Insert Figure 6 about here*****

Organism: Cognitive. Statistical equivalence was not established in relation to presence scores (SESOI = .59; $p > .05$; see Figure 7a), which precluded the acceptance of H_8. Presence in AR was conceptualized as a sense of feeling surrounded by a realistic physical/virtual environment (Georgiou & Kyza, 2017). However, a component of VR presence concerns *plausibility*, which refers to the illusion that the depicted events are really happening (Slater & Sanchez-Vives, 2016). It is possible that users' presence scores were impaired by the implausibility of experiencing dinosaurs, an extinct species, in their immediate environment. An alternative explanation is that the incomplete physical set (see Figure 4) compromised users' perception of presence.

This investigation entails one of the first attempts in tourism to employ eye-tracking in AR. The digital Tyrannosaurus Rex appeared to have captured users' visual attention effectively ($Mdn_{\text{fixations}}$ = 75%), albeit that statistical equivalence was not established with the designers' high expectations (SESOI = 8.40; $p > .05$, see Figure 7b), leading to the non-acceptance of H_9. Notwithstanding, this is a promising finding given that immersive technology allows users to navigate a scene in 360-degrees, in stark contrast to traditional modes of display (e.g., television screens; Discombe et al., 2022). Moreover, the Tyrannosaurus Rex was depicted toward the end of the AR experience, at a time when visitors are more likely to encounter satiation (i.e., reduced attention owing to repeated exposure; Rainoldi et al., 2020).

*****Insert Figure 7 about here*****

Organism: Affective. Affective valence increased from pre- to post-experience, $p < .001$, $r = .66$, in accordance with designers' intentions. This is encouraging given the high affective valence scores reported *prior* to the AR experience (*Mdn* = 7.00). Arousal scores increased from pre- to post-experience, contrary to designers' intentions, which prevented the full acceptance of H_{10}. Nonetheless, these findings indicate that AR experiences can elicit responses from the pleasant, high-arousal quadrant of the Affect Grid (Russell et al., 1989). This supports related research concerning AR travel guides (Kourouthanassis et al., 2015).

There was no evidence of statistical equivalence following the TOST procedure for either remembered pleasure or forecasted pleasure (SESOI = 16.80; *p*s > .05), leading to the non-acceptance of H_{11} (see Figure 8a and Figure 8b). Nevertheless, a promising finding to emerge from the present investigation concerns the high scores reported for remembered pleasure (*Mdn* = 62), with 95.84% of users appraising the experience positively (i.e., scores > 0; see Figure 8a). Researchers have recently emphasized the importance of measuring remembered pleasure (Godovykh & Tasci, 2020b). This is because decisions about future intentions are often predicated on memories. Hence, the present investigation provides some initial support that purpose-built AR experiences can be viable in the tourism domain.

*****Insert Figure 8 about here*****

Responses. Equivalence tests were non-significant for engagement scores (SESOI = .42; $p > .05$), precluding the acceptance of H_{12}. Nonetheless, the user engagement scores were moderate (*Mdn* = 3.66) and analogous to those obtained in other AR-related investigations in tourism (tom Dieck et al., 2018). These findings indicate that users were likely to engage with the subject matter following the completion of the AR experience. Visitor engagement is frequently cited as an important outcome of museums and so these findings attest to the potential of AR to bring the museum experience to life in an engaging manner (Serravalle et al., 2019).

Regarding negative responses, the TOST procedure (SESOI = 1.76) indicated statistical equivalence for physical demands (p = .022) and task complexity (p = .012) components of cognitive workload (see Figure 9). The remaining seven components did not reach statistical equivalence (ps > .05). With the exception of the distractions component, user scores were higher than those intended by the design team (see Figure 9). Many of the participants in the present investigation were unfamiliar with head-mounted AR devices and this could help explain the high scores observed in the task control component, which refers to the ease at which the task can be navigated (Harris, Wilson, et al., 2020).

*****Insert Figure 9 about here*****

General Discussion

AR design and user experience are often described as separate streams of research in tourism (Jingen Liang & Elliot, 2021) and yet focusing on either perspective in isolation only provides a partial view of the visitor experience (Dube et al., 2015; Ponsignon et al., 2017). Accordingly, this multi-study investigation makes an important methodological contribution to the extant literature by assessing the extent to which users' realized experience aligned with the designers' intended experience. An important theoretical contribution of the work concerns the development of a tourism-specific SOR model (see Figure 1) that depicts the relationships between an AR experience, users' internal states, and their associated behavior.

The aim of Study 1 was to explore the design intentions underlying a multi-user AR experience developed for a museum context. A mixed methods approach was adopted and the findings integrated to facilitate a comprehensive understanding of the subject matter. For example, the design intention survey revealed that the designers envisioned the AR experience to be entertaining and educational. Researchers have emphasized the importance of these constructs, often referred to as *edutainment*, in enabling successful visitor experiences in cultural contexts (Ponsignon et al., 2017). Hence, these findings provide some initial support for the applicability of AR in tourism.

The primary aim of Study 2 was to assess the extent to which users' realized experience aligned with the designers' intended experience. When considering this across the entirety of the SOR model, it appears that the greatest congruence was found in relation to the stimulus (see Figure 10). The designers' most pertinent realms of experience (i.e., entertainment and education) prompted statistically equivalent responses from users. Encouragingly, users also reported high scores for aesthetics and escapism, which alludes to the quality of the AR experience (Pine & Gilmore, 1998).

Further inspection of the findings reveals that intended and realized experience began to significantly diverge at the organism and response levels. Collectively, the findings support the notion that the design team were adept at creating an AR experience and had a sound knowledge of how users would categorize such content. However, it appears that there is an opportunity for designers to refine their expectations of how AR experiences impact users at the cognitive, affective, and behavioral level. Some disparities between designers' intended and users' realized experience are to be expected, given the recency of AR head-mounted displays (Jingen Liang & Elliot, 2021). Nonetheless, the volume of discrepancies speaks to the value of examining intended and realized experience concurrently, while placing particular emphasis on how experiences affect individuals at the organism and response levels (Scott & Le, 2017). Designers might consider the use of A/B testing to reduce any major disparities between intended and realized experience (King et al., 2017).

*****Insert Figure 10 about here*****

Theoretical Implications

AR is increasingly being implemented in tourism contexts (Loureiro et al., 2020). However, there is a paucity of research examining the factors that impact visitor behavior in relation to immersive technology use (Kim et al., 2020). Accordingly, a secondary aim of Study 2 was to examine the relationships between users' internal states and their associated behavior in relation to the AR experience. The findings provide scholars with several theoretical insights. For example, we found positive relationships between visitor engagement and presence, remembered pleasure, and forecasted pleasure (see Figure 5). These findings support the theoretical predictions of SOR models (Jacoby, 2002; Suh & Prophet, 2018), which hold that cognitive and affective states are associated with behavioral responses.

An interesting theoretical insight to emerge from the present investigation is that affective variables (i.e., remembered pleasure, forecasted pleasure) were more strongly associated with positive responses (i.e., visitor engagement) when compared to cognitive variables (i.e., presence, visual attention). However, it is noteworthy that the cross-sectional design employed herein precludes any claim of causality. Therefore, an alternative explanation for the present findings is that greater visitor engagement leads individuals to derive more positive affective responses from AR experiences.

Suh and Prophet (2018) explained that a range of individual differences can moderate the effects of immersive technology use at the organism and response levels. We sought to examine the effects of shared experiences against individual experiences given the prominent role that social interaction is theorized to play in tourism (Chen et al., 2020; Wei et al., 2019). The differences were negligible and statistically non-significant, but it is noteworthy that shared experiences constituted groups of two or three individuals. Hence, it is plausible that the small group sizes were insufficient to produce the hypothesized differences when compared to individual users. AR experiences are likely to cater for larger groups of simultaneous users as this form of technology becomes increasingly accessible.

Managerial Implications

The present investigation yields several implications for tourism managers. Researchers have previously demonstrated how AR can enhance the visitor experience by providing additional information about existing artefacts (Jung et al., 2016). The AR content in such work is typically presented to the visitor via smartphone technology which, despite being highly accessible, can be a rather solitary experience. The current findings advance this nascent line of inquiry by providing support for the efficacy of purpose-built, multi-user AR experiences that are presented through head-mounted displays. Support is evidenced by a wealth of positive user ratings across the employed SOR model (see Figure 10).

The findings revealed high user scores across each of Pine and Gilmore's (1998) four realms of experience. Objective eye-tracking data derived from the AR head-mounted display provides some initial encouragement that digital assets can be used to good effect in sustaining visitors' visual attention. Affective responses were also positive, which increases the likelihood of users sharing their experiences on social networks (Serravalle et al., 2019). Finally, users reported high scores for engagement, which is vital for the sustainability of tourism sites such as museums (Barron & Leask, 2017).

Findings from the focus group in Study 1 shine a light on how tourism managers can work in collaboration with AR designers to produce authentic experiences. The designers capitalized on validated information provided by a museum to present a scientifically accurate representation of dinosaurs in AR. This is important, as researchers have suggested that a "known facts" approach is essential for visitor education (Bec et al., 2019; Mura et al., 2017). The qualitative findings also speak to some of the challenges that designers are likely to encounter when developing similar experiences for tourism contexts. Foremost amongst these, are catering for visitors with varied levels of technical expertise and allowing sufficient time for on/offboarding.

We have also shown that purpose-built AR experiences can be effective with very little in the way of a physical set (see Figure 4). Pragmatically, this means that tourism managers can host engaging experiences without the need for visitors to gather around delicate artefacts, thereby reducing overcrowding, which has been associated with several negative outcomes (e.g., safety concerns; Yu & Egger, 2021).

Tourism managers are routinely encouraged to be receptive to new approaches pertaining to the design and analysis of visitor experiences. For example, Stienmetz et al. (2021) suggested that summary evaluations alone (e.g., SERVQUAL; Parasuraman et al., 1988) are insufficient when managing and designing tourism experiences. Consequently, there is a need to complement such measures with those that capture the sequence of events as they unfold over time (Stienmetz et al., 2021). Accordingly, it is hoped that tourism managers might employ a similar approach to that conducted in Study 2, which entailed user response data pre- (e.g., affective responses), during (i.e., visual attention), and post-experience (e.g., visitor engagement).

Limitations and Future Directions

The sample of users who took part in the present investigation was smaller than those samples employed in other immersive technology-related studies (e.g., Kim et al., 2020). Nonetheless, the sample of users (n = 48) sits comfortably toward the higher end of studies that have incorporated eye-tracking in tourism (i.e., N = 12–63; Rainoldi & Jooss, 2020). The AR head-mounted display identified when users' gaze intersected with digital assets. A logical extension to the present investigation would entail an exploration of users' gaze behavior in relation to the digital, physical, and social elements of an AR experience (Bolton et al., 2018). The intended experience could be measured from the perspectives of the service provider (e.g., visitor experience officers) in addition to those responsible for AR development. Equivalence tests were employed to determine whether users' realized experience aligned with designers' intended experience. However, researchers might explore whether users' experience exceeds designers' intentions through superiority tests.

We captured designers' intended affective journey on a scene-by-scene basis (see Figure 3). However, we refrained from collecting users' affective responses during the AR experience, as completing even single-item subjective measures would have impaired the associated eye-tracking data. Researchers might seek to employ objective measures of affective phenomena, such as skin conductance or electromyography. Combining such measures with eye-tracking would allow researchers to objectively assess the cognitive and affective states of users while they seamlessly engage with AR technology.

A cross-sectional research design was used in relation to the user testing and therefore the findings should be viewed within the frame of association. Researchers are encouraged to employ experimental designs to further the study of AR in tourism, as this would permit claims of causality. For example, assessing the visitor experience with and without such immersive technology would offer a useful addition to this line of research. The observed remembered/forecasted pleasure scores provided some initial support that users are likely to re-engage with AR technology (Karl et al., 2021). Nonetheless, longitudinal research is required to comprehensively assess the sustainability of AR technology. This research might take place beyond a museum context to scrutinize the generalizability of the present findings. Finally, we arranged constructs of interest in alignment with a SOR model (Figure 1). Such frameworks hold value to site managers and researchers alike. However, future work might expand the framework employed herein and consider the development of valid and reliable inventories that can help assess purpose-built AR experiences in tourism.

Conclusion

The present investigation contributes toward an emerging corpus of AR-related research in tourism (Loureiro et al., 2020). AR design and user experience research are often conducted separately (Jingen Liang & Elliot, 2021), despite there being considerable value in identifying the gaps between designers' intentions and users' realized experiences (Ponsignon et al., 2017). Accordingly, data were collected from designers using mixed methods, which allowed them an opportunity to convey their intent for the AR experience. Subsequently, this data served as a foundation upon which to compare users' experiences against (Smit et al., 2021). This represents a more nuanced approach to visitor experience evaluation, which typically entails comparisons between visitors' expectations and their associated experiences (Ponsignon et al., 2017). Disparities were observed between designers' intended and users' realized experience, particularly at the organism and response levels. Nonetheless, the present findings are encouraging given the infancy of AR head-mounted displays and provide a vista for tourism offerings in the envisioned metaverse.

References

Barron, P., & Leask, A. (2017). Visitor engagement at museums: Generation Y and 'Lates' events at the National Museum of Scotland. *Museum Management and Curatorship, 32*(5), 473–490. https://doi.org/10.1080/09647775.2017.1367259

Bec, A., Moyle, B., Timms, K., Schaffer, V., Skavronskaya, L., & Little, C. (2019). Management of immersive heritage tourism experiences: A conceptual model. *Tourism Management*, *72*, 117–120. https://doi.org/10.1016/J.TOURMAN.2018.10.033

Bolton, R. N., McColl-Kennedy, J. R., Cheung, L., Gallan, A., Orsingher, C., Witell, L., & Zaki, M. (2018). Customer experience challenges: Bringing together digital, physical and social realms. *Journal of Service Management*, *29*, 776–808. https://doi.org/10.1108/JOSM-04-2018-0113

Bonfanti, A., Vigolo, V., & Yfantidou, G. (2021). The impact of the Covid-19 pandemic on customer experience design: The hotel managers' perspective. *International Journal of Hospitality Management*, *94*, 102871. https://doi.org/10.1016/J.IJHM.2021.102871

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706QP063OA

Chen, X., Cheng, Z., & Kim, G. (2020). Make it memorable: Tourism experience, fun, recommendation and revisit intentions of Chinese outbound tourists. *Sustainability, 12*(5), 1904. https://doi.org/10.3390/SU12051904

Chung, N., Han, H., & Joun, Y. (2015). Tourists' intention to visit a destination: The role of augmented reality (AR) application for a heritage site. *Computers in Human Behavior*, *50*, 588–599. https://doi.org/10.1016/J.CHB.2015.02.068

Coolican, H. (2018). *Research methods and statistics in psychology* (7th ed.). Routledge.

Creswell, J. W. (2022). *A concise introduction to mixed methods research* (2nd ed.). Sage.

Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Discombe, R. M., Bird, J. M., Kelly, A., Blake, R. L., Harris, D. J., & Vine, S. J. (2022). Effects of traditional and immersive video on anticipation in cricket: A temporal occlusion study. *Psychology of Sport and Exercise, 58*, 102088. https://doi.org/10.1016/J.PSYCHSPORT.2021.102088

Dube, A., Helkkula, A., & Strandvik, T. (2015). The ripple effect: Intended and unintended app experiences. In C. von Koskull & J. Gummerus (Eds.), *The Nordic School – Alternative Perspectives on Marketing and Service Management* (pp. 273–281). Hanken School of Economics.

Ekkekakis, P. (2013). *The measurement of affect, mood, and emotion: A guide for health-behavioral research*. Cambridge University Press.

Feilzer, M. Y. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research*, *4*(1), 6–16. https://doi.org/10.1177/1558689809349691

Freeman, T. (2006). 'Best practice' in focus group research: Making sense of different views. *Journal of Advanced Nursing*, *56*(5), 491–497. https://doi.org/10.1111/J.1365-2648.2006.04043.X

Fritz, F., Susperregui, A., & Linaza, M. T. (2005). Enhancing cultural tourism experiences with augmented reality technologies. *The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST*.

Georgiou, Y., & Kyza, E. A. (2017). The development and validation of the ARI questionnaire: An instrument for measuring immersion in location-based augmented reality settings. *International Journal of Human-Computer Studies, 98*, 24–37. https://doi.org/10.1016/J.IJHCS.2016.09.014

Godovykh, M., & Tasci, A. D. A. (2020a). Customer experience in tourism: A review of definitions, components, and measurements. *Tourism Management Perspectives, 35*, 100694. https://doi.org/10.1016/J.TMP.2020.100694

Godovykh, M., & Tasci, A. D. A. (2020b). Satisfaction vs experienced utility: Current issues and opportunities. *Current Issues in Tourism*, *23*(18), 2273–2282. https://doi.org/10.1080/13683500.2020.1769573

Grand View Research. (2021). *Augmented Reality Market Size*. https://www.grandviewresearch.com/industry-analysis/augmented-reality-market

Harris, D. J., Bird, J. M., Smart, P. A., Wilson, M. R., & Vine, S. J. (2020). A framework for the testing and validation of simulated environments in experimentation and training. *Frontiers in Psychology*, *11*, 605. https://doi.org/10.3389/fpsyg.2020.00605

Harris, D. J., Wilson, M. R., & Vine, S. J. (2020). Development and validation of a simulation workload measure: The simulation task load index (SIM-TLX). *Virtual Reality*, *24*, 557–566. https://doi.org/10.1007/s10055-019-00422-9

He, Z., Wu, L., & Li, X. (2018). When art meets tech: The role of augmented reality in enhancing museum experiences and purchase intentions. *Tourism Management*, *68*, 127–139. https://doi.org/10.1016/J.TOURMAN.2018.03.003

Holmqvist, K., & Andersson, R. (2017). *Eye tracking: A comprehensive guide to methods, paradigms, and measures* (2nd ed.). Lund Eye-Tracking Research Institute.

Ingram, K. M., Espelage, D. L., Merrin, G. J., Valido, A., Heinhorst, J., & Joyce, M. (2019). Evaluation of a virtual reality enhanced bullying prevention curriculum pilot trial. *Journal of Adolescence*, *71*, 72–83. https://doi.org/10.1016/J.ADOLESCENCE.2018.12.006

Itani, O. S., & Hollebeek, L. D. (2021). Light at the end of the tunnel: Visitors’ virtual reality (versus in-person) attraction site tour-related behavioral intentions during and post-COVID-19. *Tourism Management*, *84*, 104290. https://doi.org/10.1016/J.TOURMAN.2021.104290

Jacoby, J. (2002). Stimulus-Organism-Response reconsidered: An evolutionary step in modeling (consumer) behavior. *Journal of Consumer Psychology*, *12*(1), 51–57. https://doi.org/10.1207/S15327663JCP1201_05

Jingen Liang, L., & Elliot, S. (2021). A systematic review of augmented reality tourism research: What is now and what is next? *Tourism and Hospitality Research*, *21*(1), 15–30. https://doi.org/10.1177/1467358420941913

Jung, T. H., tom Dieck, M. C., Lee, H., & Chung, N. (2016). Effects of virtual reality and augmented reality on visitor experiences in museum. In A. Inversini & R. Schegg (Eds.), *Information and communication technologies in tourism 2016* (pp. 621–635). Springer.

Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics*, *112*, 375–405. https://doi.org/10.2307/2951240

Karl, M., Kock, F., Ritchie, B. W., & Gauss, J. (2021). Affective forecasting and travel decision-making: An investigation in times of a pandemic. *Annals of Tourism Research*, *87*, 103139. https://doi.org/10.1016/J.ANNALS.2021.103139

Kim, M. J., Lee, C.-K., & Jung, T. (2020). Exploring consumer behavior in virtual reality tourism using an extended stimulus-organism-response model. *Journal of Travel Research*, *59*(1), 69–89. https://doi.org/10.1177/0047287518818915

King, R., Churchill, E. F., & Tan, C. (2017). *Designing with data: Improving user experience with A/B testing*. O’Reilly.

Kounavis, C. D., Kasimati, A. E., & Zamani, E. D. (2012). Enhancing the tourism experience through mobile augmented reality: Challenges and prospects. *International Journal of Engineering Business Management*, *4*, 10. https://doi.org/10.5772/51644

Kourouthanassis, P., Boletsis, C., Bardaki, C., & Chasanidou, D. (2015). Tourists responses to mobile augmented reality travel guides: The role of emotions on adoption behavior. *Pervasive and Mobile Computing*, *18*, 71–87. https://doi.org/10.1016/J.PMCJ.2014.08.009

Krueger, R. A., & Casey, M. A. (2014). *Focus groups: A practical guide for applied research*. SAGE Publications.

Lakens, D. (2022). Sample size justification. *Collabra: Psychology*, *8*(1), 33267. https://doi.org/10.1525/collabra.33267

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. https://doi.org/10.1177/2515245918770963

Le, D., Hadinejad, A., Moyle, B., Ma, J., & Scott, N. (2020). A review of eye-tracking methods in tourism research. In M. Rainoldi & M. Jooss (Eds.), *Eye Tracking in Tourism* (pp. 13–27). Springer.

Lishner, D. A., Cooter, A. B., & Zald, D. H. (2008). Addressing measurement limitations in affective rating scales: Development of an empirical valence scale. *Cognition and Emotion*, *22*(1), 180–192. https://doi.org/10.1080/02699930701319139

Loureiro, S. M. C., Guerreiro, J., & Ali, F. (2020). 20 years of research on virtual reality and augmented reality in tourism context: A text-mining approach. *Tourism Management*, *77*, 104028. https://doi.org/10.1016/j.tourman.2019.104028

Maguire, M., & Delahunt, B. (2017). Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *All Ireland Journal of Higher Education, 9*(3).

McCabe, S., Li, C., & Chen, Z. (2016). Time for a radical reappraisal of tourist decision making? Toward a new conceptual model. *Journal of Travel Research*, *55*(1), 3–15. https://doi.org/10.1177/0047287515592973

Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. MIT Press.

Mura, P., Tavakoli, R., & Pahlevan Sharif, S. (2017). 'Authentic but not too much': Exploring perceptions of authenticity of virtual tourism. *Information Technology & Tourism*, *17*(2), 145–159. https://doi.org/10.1007/S40558-016-0059-Y

Nevill, A., & Lane, A. M. (2007). Why self-report "Likert" scale data should not be log-transformed. *Journal of Sports Sciences*, *25*(1), 1–2. https://doi.org/10.1080/02640410601111183

Packer, J., & Ballantyne, R. (2016). Conceptualizing the visitor experience: A review of literature and development of a multifaceted model. *Visitor Studies*, *19*(2), 128–143. https://doi.org/10.1080/10645578.2016.1144023

Parasuraman, A., Zeithaml, V., & Berry, L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, *64*(1), 12–40.

Park, H. S., Seo, B.-K., Lee, G. A., & Billinghurst, M. (2020). Exploring the value of shared experience in augmented reality games. *Communications in Computer and Information Science*, *1224*, 462–469. https://doi.org/10.1007/978-3-030-50726-8_61

Park, S., & Stangl, B. (2020). Augmented reality experiences and sensation seeking. *Tourism Management*, *77*, 104023. https://doi.org/10.1016/J.TOURMAN.2019.104023

Pearce, P. L., & Packer, J. (2013). Minds on the move: New links from psychology to tourism. *Annals of Tourism Research*, *40*, 386–411. https://doi.org/10.1016/J.ANNALS.2012.10.002

Pine, B. J., & Gilmore, J. H. (1998). Welcome to the experience economy. *Harvard Business Review*, *76*, 97–105.

Ponsignon, F., & Derbaix, M. (2020). The impact of interactive technologies on the social experience: An empirical study in a cultural tourism context. *Tourism Management Perspectives*, *35*, 100723. https://doi.org/10.1016/J.TMP.2020.100723

Ponsignon, F., Durrieu, F., & Bouzdine-Chameeva, T. (2017). Customer experience design: A case study in the cultural sector. *Journal of Service Management*, *28*(4), 763–787. https://doi.org/10.1108/JOSM-01-2017-0016

Rainoldi, M., & Jooss, M. (2020). Introduction to eye tracking in tourism. In M. Rainoldi & M. Jooss (Eds.), *Eye Tracking in Tourism* (pp. 1–9). Springer.

Rainoldi, M., Yu, C.-E., & Neuhofer, B. (2020). The museum learning experience through the visitors' eyes: An eye tracking exploration of the physical context. In M. Rainoldi & M. Jooss (Eds.), *Eye Tracking in Tourism* (pp. 183–199). Springer. https://doi.org/10.1007/978-3-030-49709-5_12

Russell, J. A., & Feldman Barrett, L. (2009). Core affect. In D. Sander & K. R. Scherer (Eds.), *The Oxford companion to emotion and the affective sciences* (p. 104). Oxford University Press.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect Grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*, *57*(3), 493–502. https://doi.org/10.1037/0022-3514.57.3.493

Scott, N., & Le, D. (2017). Tourism experience: A review. In N. Scott, J. Gao, & J. Ma (Eds.), *Visitor Experience Design* (pp. 30–52). CABI.

Scott, N., Zhang, R., Le, D., & Moyle, B. (2019). A review of eye-tracking research in tourism. *Current Issues in Tourism*, *22*(10), 1244–1261. https://doi.org/10.1080/13683500.2017.1367367

Serravalle, F., Ferraris, A., Vrontis, D., Thrassou, A., & Christofi, M. (2019). Augmented reality in the tourism industry: A multi-stakeholder analysis of museums. *Tourism Management Perspectives*, *32*, 100549. https://doi.org/10.1016/J.TMP.2019.07.002

Skavronskaya, L., Scott, N., Moyle, B., Le, D., Hadinejad, A., Zhang, R., Gardiner, S., Coghlan, A., & Shakeela, A. (2017). Cognitive psychology and tourism research: State of the art. *Tourism Review*, *72*(2), 221–237.

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI*, *3*, 74. https://doi.org/10.3389/frobt.2016.00074

Smink, A. R., van Reijmersdal, E. A., van Noort, G., & Neijens, P. C. (2020). Shopping in augmented reality: The effects of spatial presence, personalization and intrusiveness on app and brand responses. *Journal of Business Research*, *118*, 474–485. https://doi.org/10.1016/J.JBUSRES.2020.07.018

Smit, B., Melissen, F., Font, X., & Gkritzali, A. (2021). Designing for experiences: A meta-ethnographic synthesis. *Current Issues in Tourism*, *24*(21), 2971–2989. https://doi.org/10.1080/13683500.2020.1855127

Stienmetz, J., Kim, J., Xiang, Z., & Fesenmaier, D. R. (2021). Managing the structure of tourism experiences: Foundations for tourism design. *Journal of Destination Marketing & Management*, *19*, 100408. https://doi.org/10.1016/J.JDMM.2019.100408

Suh, A., & Prophet, J. (2018). The state of immersive technology research: A literature analysis. *Computers in Human Behavior*, *86*, 77–90. https://doi.org/10.1016/J.CHB.2018.04.019

Sweller, J. (1999). *Instructional design in technical areas.* ACER Press.

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson Education.

Taheri, B., Jafari, A., & O'Gorman, K. (2014). Keeping your audience: Presenting a visitor engagement scale. *Tourism Management, 42*, 321–329. https://doi.org/10.1016/J.TOURMAN.2013.12.011

Terry, G., Hayfield, N., Clarke, V., & Braun, V. (2017). Thematic Analysis. In C. Willig & W. Stainton-Rogers (Eds.), *The SAGE Handbook of Qualitative Research in Psychology* (pp. 17–37). SAGE Publications.

tom Dieck, M. C., Jung, T. H., & Rauschnabel, P. A. (2018). Determining visitor engagement through augmented reality at science festivals: An experience economy perspective. *Computers in Human Behavior, 82,* 44–53. https://doi.org/10.1016/J.CHB.2017.12.043

Trunfio, M., & Campana, S. (2020). A visitors' experience model for mixed reality in the museum. *Current Issues in Tourism, 23,* 1053–1058. https://doi.org/10.1080/13683500.2019.1586847

Tucker, H., & Shelton, E. J. (2018). Tourism, mood and affect: Narratives of loss and hope. *Annals of Tourism Research, 70,* 66–75. https://doi.org/10.1016/J.ANNALS.2018.03.001

Tussyadiah, I. P. (2014). Toward a theoretical foundation for experience design in tourism. *Journal of Travel Research, 53*(5), 543–564. https://doi.org/10.1177/0047287513513172

Voss, C., Roth, A. V., & Chase, R. B. (2008). Experience, service operations strategy, and services as destinations: Foundations and exploratory investigation. *Production and Operations Management, 17*(3), 247–266. https://doi.org/10.3401/poms.1080.0030

Walls, A. R., Okumus, F., & Wang, Y. (2011). Cognition and affect interplay: A framework for the tourist vacation decision-making process. *Journal of Travel & Tourism Marketing*, *28*(5), 567–582. https://doi.org/10.1080/10548408.2011.588121

Wattanacharoensil, W., & La-ornual, D. (2019). A systematic review of cognitive biases in tourist decisions. *Tourism Management*, *75*, 353–369. https://doi.org/10.1016/J.TOURMAN.2019.06.006

Wei, C., Zhao, W., Zhang, C., & Huang, K. (2019). Psychological factors affecting memorable tourism experiences. *Asia Pacific Journal of Tourism Research*, *24*(7), 619–632. https://doi.org/10.1080/10941665.2019.1611611

WTTC. (2021). *Travel & Tourism Economic Impact 2021*. https://wttc.org/Portals/0/Documents/Reports/2021/Global%20Economic%20Impact%20and%20Trends%202021.pdf?ver=2021-07-01-114957-177

Yeh, C.-H., Wang, Y.-S., Li, H.-T., & Lin, S.-Y. (2017). The effect of information presentation modes on tourists' responses in Internet marketing: The moderating role of emotions. *Journal of Travel & Tourism Marketing*, *34*(8), 1018–1032.

Yu, J., & Egger, R. (2021). Tourist experiences at overcrowded attractions: A text analytics approach. *Information and Communication Technologies in Tourism 2021*, 231–243. https://doi.org/10.1007/978-3-030-65785-7_21

Zenko, Z., Ekkekakis, P., & Ariely, D. (2016). Can you have your vigorous exercise and enjoy it too? Ramping intensity down increases postexercise, remembered, and forecasted pleasure. *Journal of Sport & Exercise Psychology*, *38*(2), 149–159. https://doi.org/10.1123/jsep.2015-0286

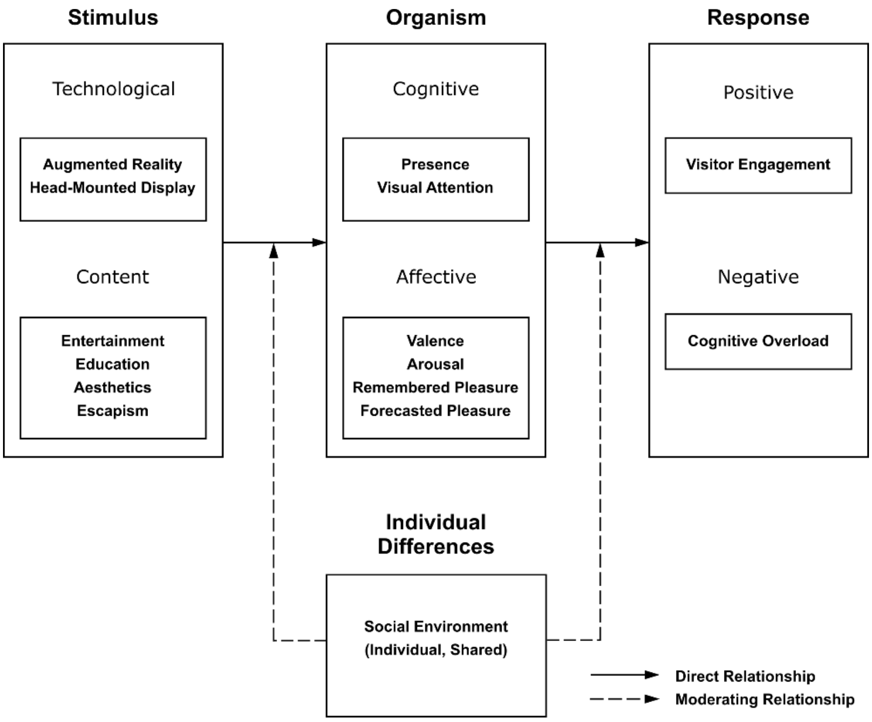

Figure 1. The Stimulus-Organism-Response model employed in the present investigation.

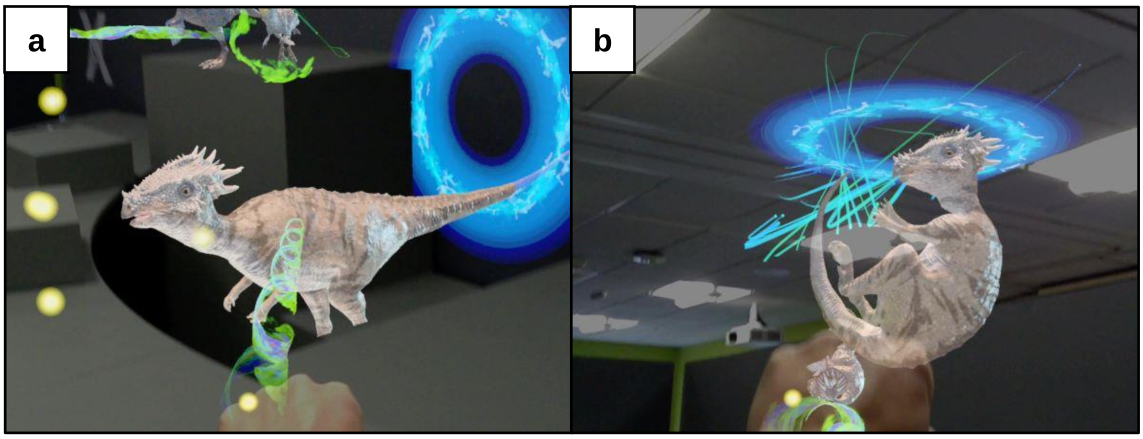

Figure 2. Spotting a ghost dinosaur (a) and returning it to their dimension (b).

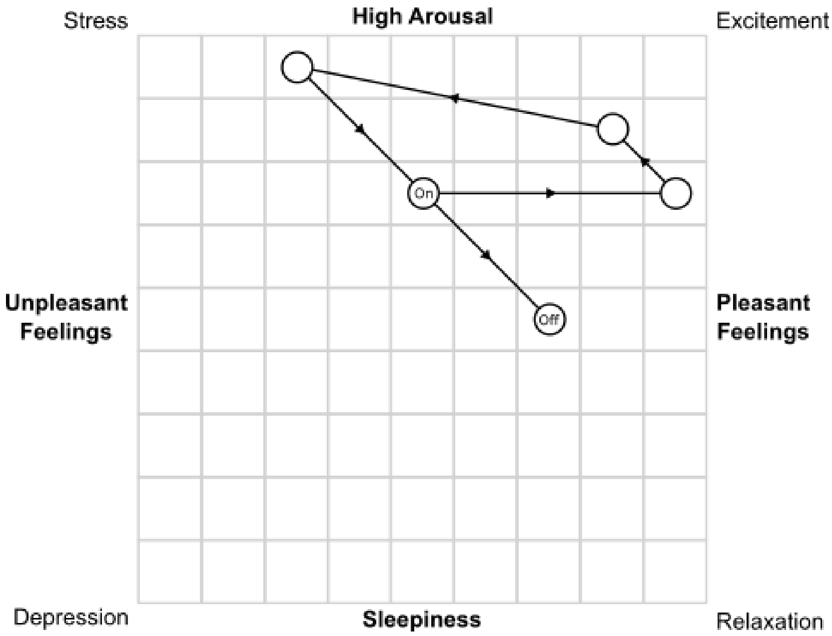

Figure 3. The affective journey associated with the augmented reality experience.

Note. On = Onboarding or pre-experience, Off = Offboarding or post-experience.

Figure 4. The physical set associated with the augmented reality experience.

Note. Users were required to stand behind one of the six colored squares placed along the perimeter of the desk.

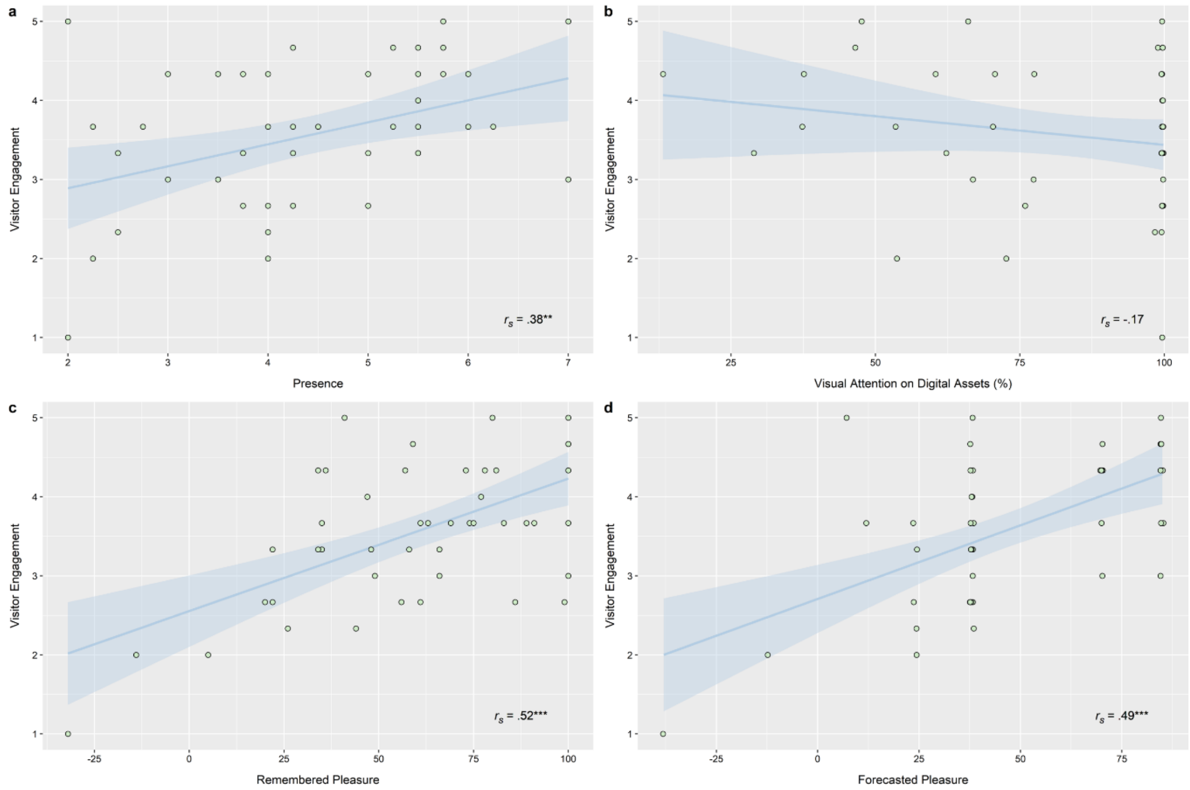

Figure 5. Correlation plots between visitor engagement and presence (a), visual attention (b), remembered pleasure (c), and forecasted pleasure (d). *Note.* ***p* < .01. ****p* < .001.

Figure 6. Raincloud plot depicting designer and user scores for each realm of experience proposed by Pine and Gilmore (1998). *Note.* Each density plot displays the median and interquartile range. Ex. = experience. $^{**}p < .01$.

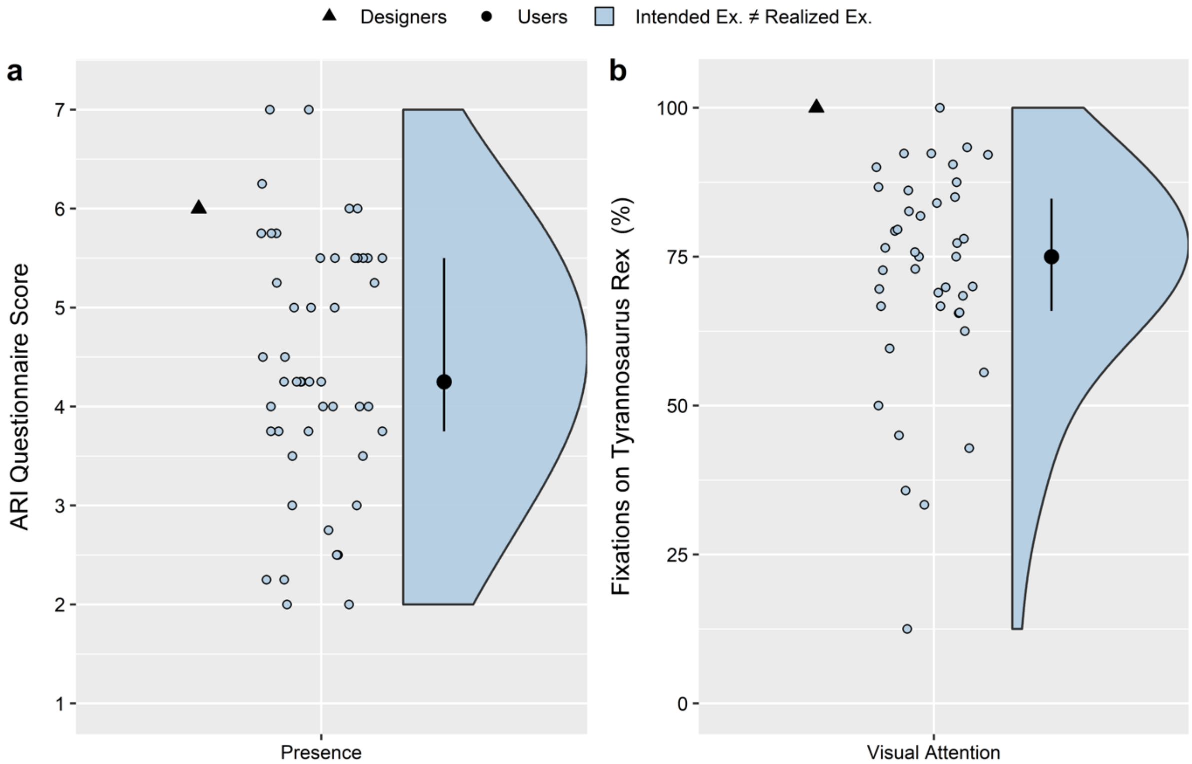

Figure 7. Raincloud plots depicting designer and user scores for presence (a) and visual attention (b). *Note.* Each density plot displays the median and interquartile range. Ex. = experience; ARI = Augmented Reality Immersion.

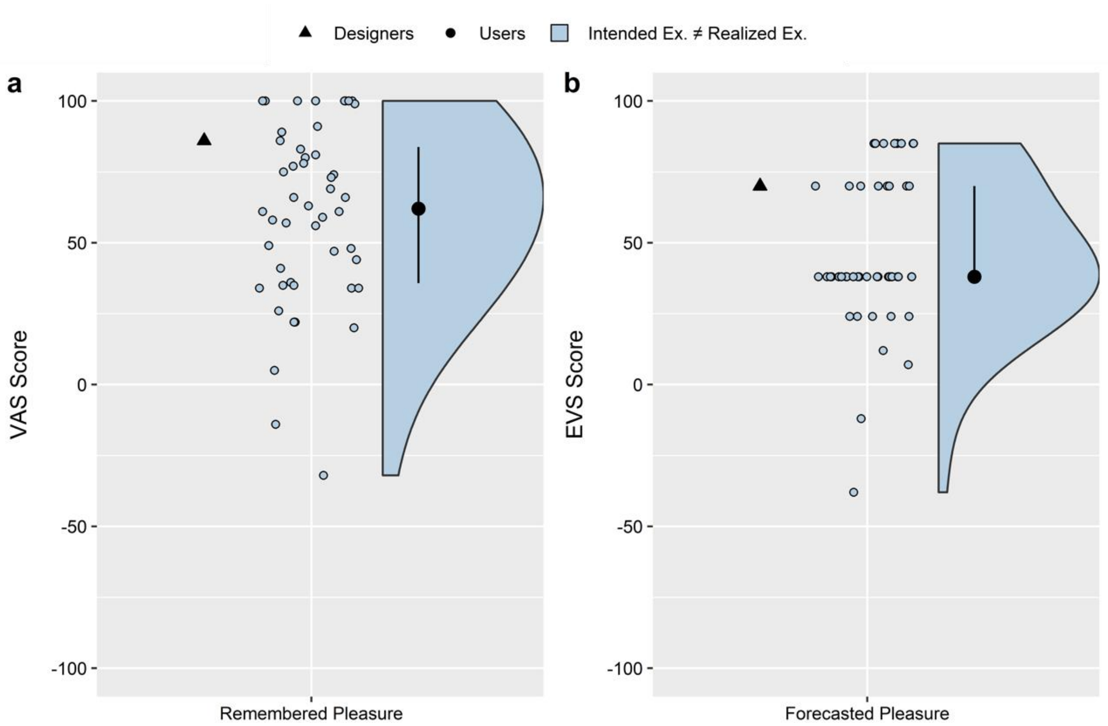

Figure 8. Raincloud plots depicting user and designer scores for remembered pleasure (a) and forecasted pleasure (b). *Note.* Each density plot displays the median and interquartile range. Ex. = experience; VAS = Visual Analogue Scale; EVS = Empirical Valence Scale.

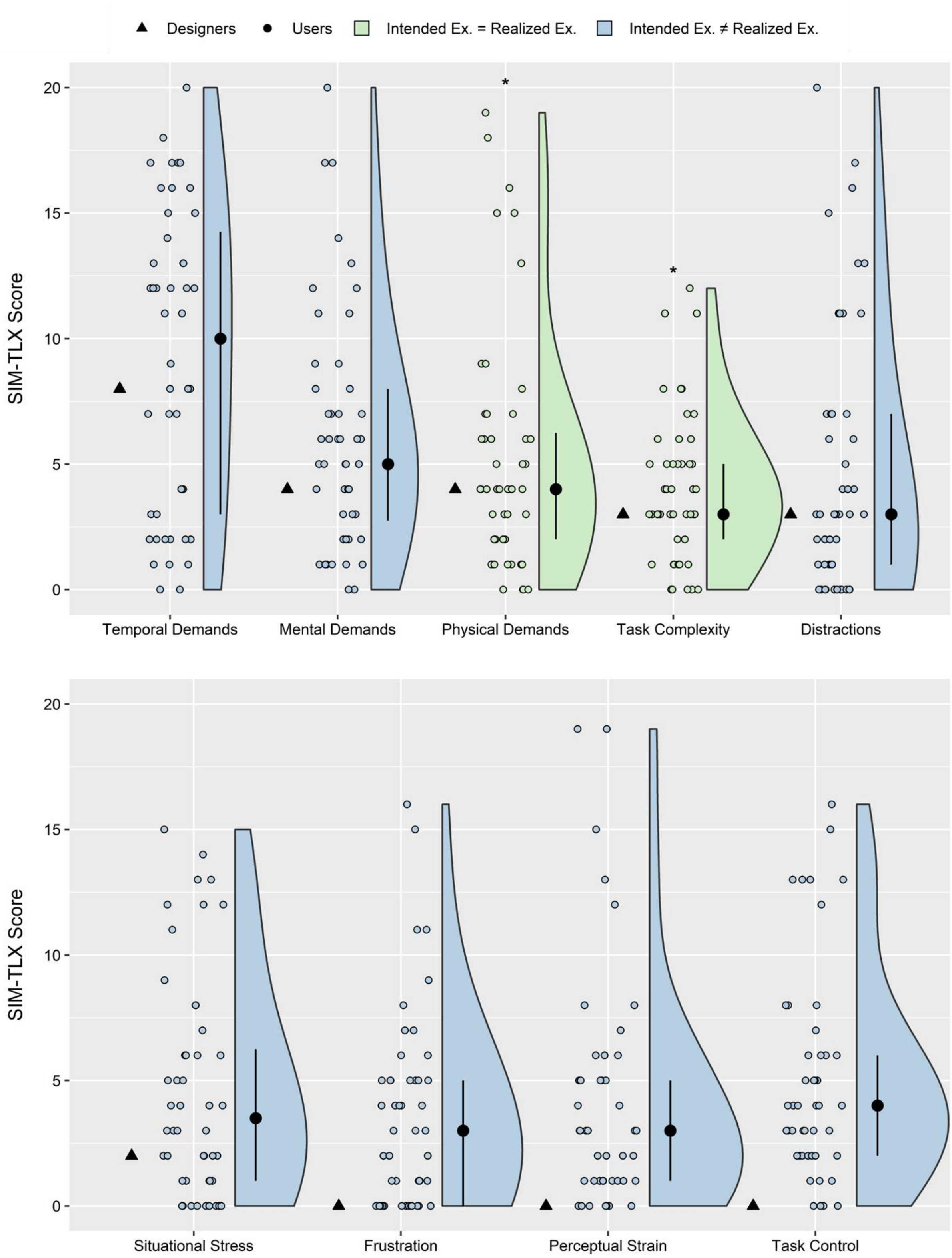

Figure 9. Raincloud plots depicting designer and user scores for cognitive workload. *Note.* Each density plot displays the median and interquartile range. Ex. = experience; SIM-TLX = Simulation Task Load Index. *$p < .05$.

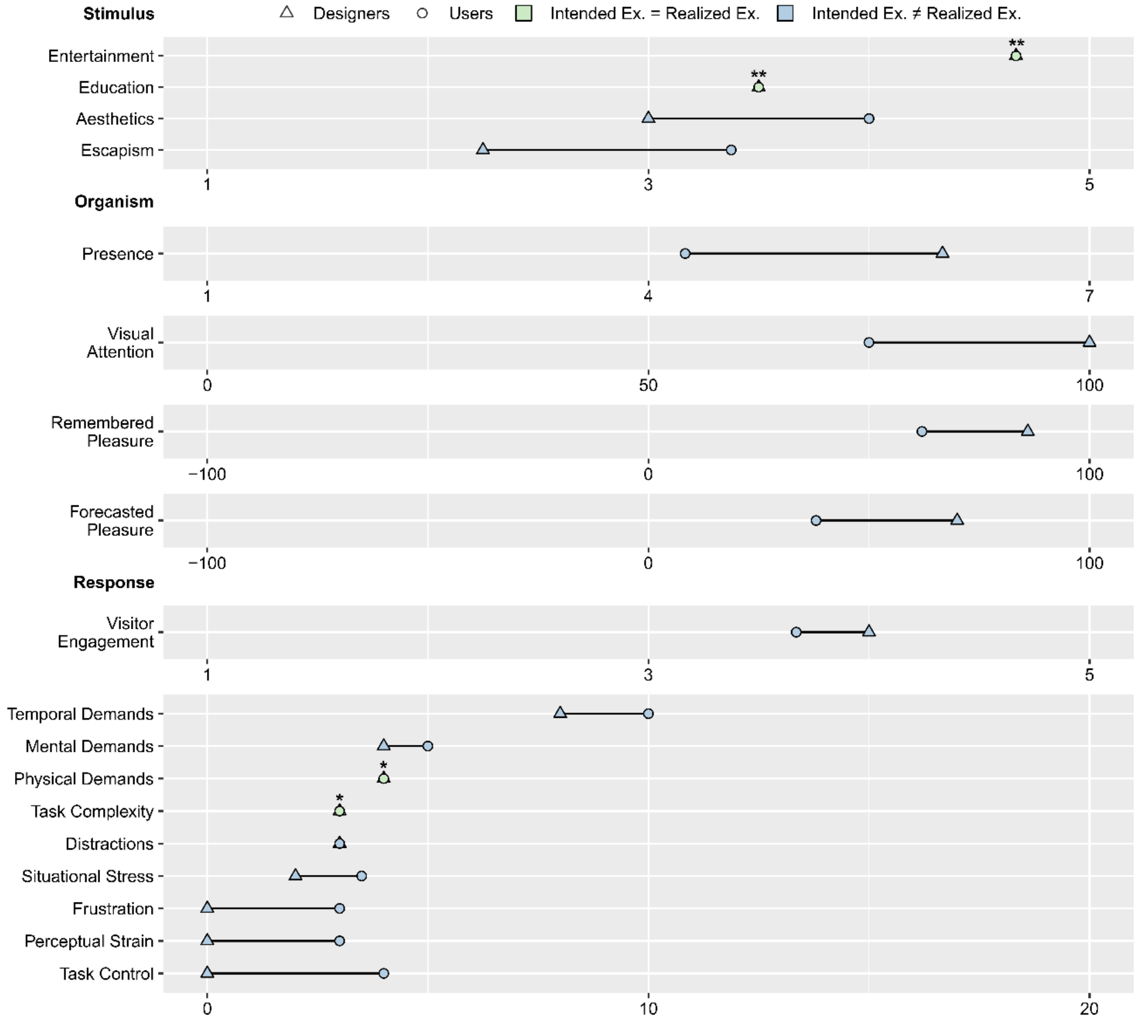

Figure 10. Dumbbell plot depicting designer and user median scores for each construct across the Stimulus-Organism-Response model. *Note*. Ex. = experience. *$p < .05$. **$p < .01$.

Supplementary Material 1

Pragmatism

The present investigation was conducted in alignment with a pragmatic research paradigm (Feilzer, 2010). This position directs scholars' attention toward research questions and methodologies that can make the greatest applied impact to the individuals and groups that they examine (Morgan, 2014). Pragmatism does not prioritize metaphysical concerns such as the nature of reality and the possibility of an objective truth (Morgan, 2014). However, the pragmatist position discards the assumptions of an objective reality, and that one interpretation can more accurately represent the truth than another. Mixed methods complements pragmatism (Mitchell, 2018). This is because mixed methods allow the unique perspectives of participants to emerge via qualitative methods, while also allowing an examination of the relationships between measurable variables via quantitative methods (Feilzer, 2010).

References

Feilzer, M. Y. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research, 4*(1), 6–16. https://doi.org/10.1177/1558689809349691

Mitchell, A. (2018). A review of mixed methods, pragmatism and abduction techniques. *Electronic Journal of Business Research Methods, 16*(3), 103–116.

Morgan, D. L. (2014). Pragmatism as a paradigm for social research. *Qualitative Inquiry, 20*(8), 1045–1053. https://doi.org/10.1177/1077800413513733

Supplementary Material 2

Table 1

User Experience Survey Items

SOR Construct	Item	Source
Stimulus		
Content Stimuli	Entertainment 1: The AR experience was amusing.	tom Dieck et al. (2018)
	Entertainment 2: The AR experience was entertaining.	
	Entertainment 3: The AR experience was fun.	
	Education 1: I learned something new during the AR experience.	
	Education 2: The experience made me more knowledgeable.	
	Education 3: It stimulated my curiosity to learn new things.	
	Education 4: It was a real learning experience.	
	Aesthetics 1: The setting of the AR experience was very attractive.	
	Aesthetics 2: The AR experience was very pleasant.	
	Aesthetics 3: I felt a real sense of harmony.	
	Escapism 1: I felt I played a different character when using the AR application.	
	Escapism 2: I felt like I was living in a different time or place.	
	Escapism 3: The AR experience let me imagine being someone else.	
	Escapism 4: I completely escaped from reality.	
Organism		
Cognitive	Presence 1: The activity felt so authentic that it made me think that the virtual characters/objects existed for real.	Georgiou and Kyza (2017)
	Presence 2: I felt that what I was experiencing was something real, instead of a fictional activity.	
	Presence 3: I was so involved in the activity, that in some cases I wanted to interact with the virtual characters/objects directly.	
	Presence 4: I so was involved, that I felt that my actions could affect the activity.	

Continued

Table 1 Continued

SOR Construct	Item	Source
Affective	Remembered Pleasure: How did the experience make you feel? Forecasted Pleasure: If you repeated the experience again, how do you think it would make you feel?	Lishner et al. (2008); Zenko et al. (2016)
Response		
Positive	Visitor Engagement 1: This experience has motivated me to find out more about the history of dinosaurs. Visitor Engagement 2: This experience has motivated me to find out more about dinosaur research. Visitor Engagement 3: This experience has motivated me to participate in additional AR activities.	tom Dieck et al. (2018)
Negative	Mental Demands: How mentally fatiguing was the task? Physical Demands: How physically fatiguing was the task? Temporal Demands: How hurried or rushed did you feel during the task? Frustration: How insecure, discouraged, irritated, stressed or annoyed were you? Task Complexity: How complex was the task? Situational Stress: How stressed did you feel while performing the task? Distractions: How distracting was the task environment? Perceptual Strain: How uncomfortable/irritating were the visual and auditory aspects of the task? Task Control: How difficult was the task to control/navigate?	Harris et al. (2020)

Note. The Affect Grid (Russell et al., 1989) was also administered pre- and post-experience. SOR = Stimulus-Organism-Response; AR = augmented reality.

References

Georgiou, Y., & Kyza, E. A. (2017). The development and validation of the ARI questionnaire: An instrument for measuring immersion in location-based augmented reality settings. *International Journal of Human-Computer Studies, 98*, 24–37. https://doi.org/10.1016/j.ijhcs.2016.09.014

Harris, D. J., Wilson, M. R., & Vine, S. J. (2020). Development and validation of a simulation workload measure: The simulation task load index (SIM-TLX). *Virtual Reality*, *24*, 557–566. https://doi.org/10.1007/s10055-019-00422-9

Lishner, D. A., Cooter, A. B., & Zald, D. H. (2008). Addressing measurement limitations in affective rating scales: Development of an empirical valence scale. *Cognition and Emotion*, *22*(1), 180–192. https://doi.org/10.1080/02699930701319139

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect Grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology, 57*(3), 493–502. https://doi.org/10.1037/0022-3514.57.3.493

tom Dieck, M. C., Jung, T. H., & Rauschnabel, P. A. (2018). Determining visitor engagement through augmented reality at science festivals: An experience economy perspective. *Computers in Human Behavior, 82,* 44–53. https://doi.org/10.1016/j.chb.2017.12.043

Zenko, Z., Ekkekakis, P., & Ariely, D. (2016). Can you have your vigorous exercise and enjoy it too? Ramping intensity down increases postexercise, remembered, and forecasted pleasure. *Journal of Sport & Exercise Psychology, 38*(2), 149–159. https://doi.org/10.1123/jsep.2015-0286

Supplementary Material 3

Data Analysis

Objective eye-tracking data was analyzed using MATLAB R2018a (MathsWorks; Natick, MA, USA). The augmented reality (AR) head-mounted displays provided 'gaze-in-world' coordinates in three dimensions (i.e., *x*, *y*, *z*), which indicated the point where gaze converged on either a physical or digital asset. The data file was imported into MATLAB and the gaze location coordinates were initially passed through a three-frame median filter and de-noised using a second-order, zero-lag Butterworth filter with a 15 Hz cut-off (Cesqui et al., 2015; Fooken & Spering, 2020).

Fixations were identified using a spatial dispersion algorithm, which was adapted from the EyeMMV toolbox for MATLAB (Krassanakis et al., 2014). The spatial dispersion algorithm identifies periods of stable visual fixations by grouping successive gaze points into clusters, based on their spatial similarity. A minimum duration criterion of 100 ms and a maximum spatial dispersion of 3° of visual angle was allocated (Salvucci & Goldberg, 2000). Coordinate positions were converted to visual angle based on the distance from the fixation location to the user. Fixations that did not pass a confidence threshold of 70% were discarded. The AR device also identified when users' gaze intersected with digital assets. Accordingly, we were able to determine the primary object of each fixation. All MATLAB scripts are available online (https://osf.io/BT3UV/).

References

Cesqui, B., Mezzetti, M., Lacquaniti, F., & D'Avella, A. (2015). Gaze behavior in one-handed catching and its relation with interceptive performance: What the eyes can't tell. *PLoS ONE, 10*(3), e0119445. https://doi.org/10.1371/journal.pone.0119445

Fooken, X. J., & Spering, M. (2020). Eye movements as a readout of sensorimotor decision processes. *Journal of Neurophysiology*, *123*(4), 1439–1447. https://doi.org/10.1152/jn.00622.2019

Krassanakis, V., Filippakopoulou, V., & Nakos, B. (2014). EyeMMV toolbox: An eye movement post-analysis tool based on a two-step spatial dispersion threshold for fixation identification. *Journal of Eye Movement Research, 7*(1), 1–10.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the Symposium on Eye Tracking Research & Applications*, 71–78.

Supplementary Material 4

Table 1

User Experience Survey Results

SOR Construct	Item	*n*	*M*	*SD*	Skewness		Kurtosis	
					Statistic	Std. Error	Statistic	Std. Error
Stimulus								
Content Stimuli	Entertainment 1	48	4.67	0.60	-1.64	0.34	1.71	0.67
	Entertainment 2	48	4.38	0.82	-1.05	0.34	0.15	0.67
	Entertainment 3	48	4.33	0.72	-0.96	0.34	0.91	0.67
	Education 1	48	3.08	1.03	-0.30	0.34	-0.76	0.67
	Education 2	48	3.69	1.15	-0.74	0.34	-0.04	0.67
	Education 3	48	3.35	1.02	-0.15	0.34	-0.18	0.67
	Education 4	48	3.83	1.00	-0.73	0.34	0.20	0.67
	Aesthetics 1	48	4.10	0.86	-0.63	0.34	-0.32	0.67
	Aesthetics 2	48	4.29	0.77	-0.85	0.34	0.22	0.67
	Aesthetics 3	48	3.35	1.00	0.15	0.34	-0.42	0.67
	Escapism 1	48	3.33	1.23	-0.39	0.34	-0.90	0.67
	Escapism 2	48	3.10	1.26	-0.07	0.34	-0.88	0.67
	Escapism 3	48	3.21	1.30	-0.04	0.34	-1.26	0.67
	Escapism 4	48	3.19	1.28	-0.30	0.34	-0.84	0.67
Organism								
Cognitive	Presence 1	48	3.88	1.62	0.02	0.34	-0.44	0.67
	Presence 2	48	3.83	1.52	0.10	0.34	-0.51	0.67
	Presence 3	48	4.73	1.63	-0.37	0.34	-0.33	0.67
	Presence 4	48	5.25	1.39	-0.57	0.34	-0.36	0.67
Affective	Pre-Ex. Valence	48	6.54	1.75	-0.74	0.34	-0.10	0.67
	Pre-Ex. Arousal	48	6.04	1.88	-0.65	0.34	-0.05	0.67
	Post-Ex. Valence	48	7.81	1.20	-1.02	0.34	0.99	0.67
	Post-Ex. Arousal	48	6.88	1.79	-0.89	0.34	-0.26	0.67
	Remembered Pleasure	48	60.17	31.65	-0.71	0.34	0.37	0.67
	Forecasted Pleasure	48	45.90	26.80	-0.41	0.34	0.88	0.67

Continued

Table 1 Continued

SOR Construct	Item	*n*	*M*	*SD*	Skewness		Kurtosis	
					Statistic	Std. Error	Statistic	Std. Error
Response								
Positive	Engagement 1	48	3.38	1.06	-0.04	0.34	-0.83	0.67
	Engagement 2	48	3.13	1.08	0.06	0.34	-0.87	0.67
	Engagement 3	48	4.19	0.98	-1.38	0.34	1.73	0.67
Negative	Mental Demands	48	6.13	4.70	1.09	0.34	0.93	0.67
	Physical Demands	48	5.23	4.79	1.44	0.34	1.58	0.67
	Temporal Demands	48	9.21	5.94	-0.07	0.34	-1.35	0.67
	Frustration	48	3.58	3.92	1.41	0.34	2.03	0.67
	Task Complexity	48	3.98	2.92	0.88	0.34	0.71	0.67
	Situational Stress	48	4.63	4.39	0.92	0.34	-0.24	0.67
	Distractions	48	4.81	5.28	1.24	0.34	0.66	0.67
	Perceptual Strain	48	4.04	4.67	1.86	0.34	3.36	0.67
	Task Control	48	5.02	4.14	1.19	0.34	0.68	0.67

Note. SOR = Stimulus-Organism-Response; Ex. = experience.

Supplementary Material 5

Table 1

Hypotheses Tested in the Present Investigation

	Hypothesis	Description
User Responses	H_1	Positive relationships between presence–visitor engagement and between visual attention–visitor engagement.
	H_2	Positive relationships between remembered pleasure–visitor engagement and between forecasted pleasure–visitor engagement.
	H_3	Stronger relationships between cognitive variables–visitor engagement when compared to the relationships between affective variables–visitor engagement.
	H_4	Greater visitor engagement for those who take part in a shared experience when compared to individual users.
	H_5	More positive remembered/forecasted pleasure for those who take part in a shared experience when compared to individual users.
	H_6	Lower presence and visual attention towards digital assets for those who take part in a shared experience when compared to individual users.
Intended and Realized Experience	H_7	Statistical equivalence between designers' intentions and users' realized experience for each realm of experience (i.e., entertainment, education, aesthetics, escapism).
	H_8	Statistical equivalence between designers' intentions and users' realized experience for perceptions of presence.
	H_9	Statistical equivalence between designers' intentions and users' realized experience in relation to visual attention toward the digital Tyrannosaurus Rex.
	H_{10}	Users' affective valence will increase, and arousal will decrease from pre- to post-experience, in alignment with designers' intentions.
	H_{11}	Statistical equivalence between designers' intentions and users' realized experience for remembered pleasure and forecasted pleasure.
	H_{12}	Statistical equivalence between designers' intentions and users' realized experience for visitor engagement.
	H_{13}	Statistical equivalence between designers' intentions and users' realized experience for cognitive workload, with temporal demands prompting the highest user scores.

Note. $H_{1–6}$ were derived from the extant literature. $H_{7–13}$ were developed following mixed methods data collection with the design team.

Supplementary Material 6

Table 1

Results of Wilcoxon Rank Sum Tests Between Individual and Shared Experience

Variable	Individual Ex.			Shared Ex.			*W*	*p*	*r*
	n	*Mdn*	*IQR*	*n*	*Mdn*	*IQR*			
Presence	18	5.50	1.62	30	4.12	1.31	386.5	.066	.36
Visual Attention	18	88.10	36.40	30	99.60	23.80	201.0	.585	.19
Remembered Pleasure	18	68.00	54.00	30	62.00	39.20	298.5	1.000	.09
Forecasted Pleasure	18	38.00	42.50	30	38.00	32.00	262.5	1.000	.02
Visitor Engagement	18	3.67	1.25	30	3.50	1.58	358.0	.239	.27

Note. Holm-Bonferroni corrections were applied to *p* values help control family-wise error. Ex. = experience; *Mdn* = median; *IQR* = interquartile range.