Received: 3 May 2023 | Accepted: 22 May 2024

DOI: 10.1111/cura.12637

CURATOR THE MUSEUM JOURNAL

ARTICLE

Evaluating museum exhibits: Quantifying visitor experience and museum impact with user experience methodologies

Ellie King[1,2] | **M. Paul Smith**[2] | **Paul F. Wilson**[1] | **Janet F. Stott**[2] | **Mark A. Williams**[1]

[1]Warwick Manufacturing Group, University of Warwick, Coventry, UK

[2]Oxford University Museum of Natural History, Oxford, UK

Correspondence
Ellie King, Warwick Manufacturing Group, University of Warwick, Coventry, UK.
Email: ellie.king.1@warwick.ac.uk

Funding information
Arts and Humanities Research Council

Abstract

Underpinned by the Model for Museum Exhibit User Experience (MEUX; King et al., *Visitor Studies*, 2023, 26, 59), this paper develops and presents an evaluation methodology for museum exhibits that utilizes existing methodologies from the user experience sector adapted for the museum and cultural heritage sectors. Two studies are presented: an in-depth evaluation of the *Meat the Future* exhibition at Oxford University Museum of Natural History and then a comparative study between this exhibition and two other permanent exhibits at the museum. Quantitative and qualitative data provide a nuanced picture of each exhibit from the visitor perspective and showcase the benefits of the MEUX methods of evaluation. Results show how three different exhibits are constructed in different ways, providing different visitor experiences and outcomes. They are directly compared with identify statistical differences, but do not impose a judgment as to whether any exhibit is better than another. With detailed, nuanced and rigorous data capturing visitor experiences of engaging with exhibits, the MEUX evaluation methodology allows for more sophisticated, standardized and efficient evaluation practices within the sector, with results that directly support further development of exhibits and exhibitions.

INTRODUCTION

A fundamental activity for museums and heritage sites worldwide is the development and presentation of temporary and permanent museum exhibits. These come in many different shapes and sizes, ranging from simple glass-case displays of museum objects to full sensory and immersive experiences. The key role of exhibits is for visitors to the museum to engage with them: they look at objects, read text, watch videos, take part in digital interactives, and more. A model to conceptualize this visitor experience has been produced; the Museum Exhibit User Experience (MEUX; King et al., 2023), which describes how the features, qualities, and aims of an exhibit come together, and how a visitor perceives and interacts with them.

Operating hand in hand with exhibitions is the evaluation of them. It is common for museums to seek visitor opinions of exhibits and exhibitions, including what visitors learned, their enjoyment, and if there were anything they would improve. While evaluation is widespread, it is often hampered by a lack of standardized methods, generalisable results, and quantitative methodologies, and therefore has limited ability to facilitate change within organizations (Davies & Heath, 2013; Johanson & Glow, 2015; Kubarek & Trainer, 2015). With this in mind, the MEUX Model has been developed to provide rigorous and sophisticated methods of evaluation, using methods directly taken from the User Experience industry. The MEUX evaluation methodology is explored in this paper in two ways: a detailed study of a temporary exhibition, *Meat the Future*, at Oxford University Museum of Natural History (OUMNH); then a comparison of *Meat the Future* with two different permanent exhibits at OUMNH. The paper highlights the benefits and real power of the methodology in enabling comparison of different types of exhibit using rigorous and standardized data to assess the impact of different visitor experiences. The paper discusses related work concerning the MEUX Model, museum evaluation, and User Experience evaluation methodologies, before detailing method development, results from the in-depth study of *Meat the Future*, and results from the comparative studies.

RELATED WORK

A model for museum exhibit user experience

The Model for Museum Exhibit User Experience (King et al., 2023) was developed and adapted from the User Experience sector, and specifically a model of User Experience postulated by Hassenzahl (2003). This model conceptualizes user-product interactions, and acts as a framework for understanding user perceptions of product features and product quality. The model was adapted for use in the museum sector due to a number of similarities between how users interact with products and how visitors interact with museum exhibits. Full development of the Model is explored in King et al. (2023), but key justifications for its use are cited below.

First, it acknowledges the complex identities of users that affect product interaction, which echoes the extensive literature exploring the complex identities of museum visitors (Falk & Dierking, 1992; Hein, 1998). Second, it acknowledges there is no guarantee that the outcome a designer intends for a product is actually received and engaged with by end users. This reflects an understanding in the museum literature that there is no guarantee that visitors will engage and learn from exhibits as intended (Falk & Dierking, 2016; Hein, 1998; Rennie & Johnstone, 2004), with several instances of intentions and consequences not matching up (Baldioli et al., 2022; Ji et al., 2023; Krause & Davison, 2021; Packer et al., 2022). Finally, it recognizes that product quality is judged on a range of qualities, including how the product performs in tasks, but also its more general feel, the emotional response, and the sense of enjoyment and satisfaction. Once again, this reflects the wide range of experiences of museum learning environments, which include learning, but also enjoyment, satisfaction, sense of value, and

emotional engagement (Angeli et al., 2020; Baker, 2015; Del Chiappa et al., 2014; Gardner, 2015; Hoare, 2020; Hooper-Greenhill, 1992, 2000; Luebke, 2018; Macdonald, 2006; May et al., 2022; Munro, 2014; O'Connor et al., 2020; Price et al., 2021; Price & Applebaum, 2022; Radywyl et al., 2015; Robinson, 2021; Rodehn, 2020; Roschelle, 1995; Watson, 2015).

These reasons highlight several similarities between the User Experience and Museums sectors, which provides a reasonable basis for the development of the Model of Museum Exhibit User Experience. This Model acts as a starting point for developing an evaluation tool that can be used to measure the user experience of specific exhibits. Throughout the text, Model components are italicized.

The Model for MEUX (Figure 1) conceptualizes visitor experience and interactions with a museum exhibit or exhibition from both the *Museum Perspective* and the *Visitor Perspective*. From the *Museum Perspective*, *Exhibition Features* such as text, objects, design, accessibility, and digital interactives come together to produce an *Intended Exhibition Character*, which is how the museum intends the exhibit to be perceived by visitors. This *Character* is divided into two types of *Qualities*: *Pragmatic Qualities*, which are concerned with the effectiveness of getting the exhibition message/s across to visitors; and *Hedonic Qualities*, which are about the experience of the visitor in receiving and engaging with that message. These then form *Key Aims* that can be something *For the Museum*, such as particular visitor demographics, as well as what they would like *Visitors to Know*, *Feel*, and *Do*.

From the *Visitor Perspective*, *Exhibition Features* come together to form an *Apparent Exhibition Character*, which is how the visitor perceives the exhibition through both its *Pragmatic* and *Hedonic Qualities*. This *Apparent Character* is not necessarily what the museum *Intended*, and this is due to the *Visitor Situation*: factors concerning the visitor's identity that mediates how they perceive and engage with an exhibit. This involves elements such as the visitor's age, gender, and education level, but also the different learning styles presented in the exhibit and their motivations for visiting the museum that day. This mediated *Visitor Situation* then forms the actual *Visitor Gains* of the exhibit, again in the layers of *Know*, *Feel*, and *Do*.

The presentation and exploration of the MEUX Model highlighted the detailed conceptualisation of visitor experience with exhibits, and, with a simple evaluation method, explored the relative priorities of cultural institutions and their visitors. The study found a significant

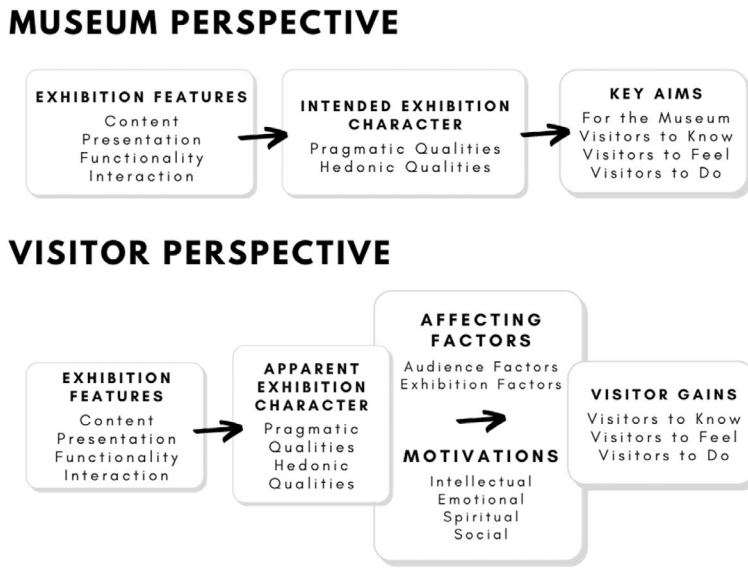


FIGURE 1 Museum Exhibit User Experience (MEUX) Model. King et al. (2023).

gulf between institutional and visitor priorities concerning exhibits, highlighting the need for further research in this area. The research team identified the development of more detailed evaluation methods as a priority, which would involve understanding the *Apparent Exhibition Character* of an exhibit and how this was mediated by a *Visitor Situation* to produce actual *Visitor Gains*. It was also a priority to understand, based on the MEUX Model, what inputs of *Exhibition Features* influenced visitor outcomes and how they perceived an exhibit.

Museum visitor learning and experience

The conceptual MEUX Model brings together extensive literature that has explored visitor learning and engagement in museums. In the theoretical literature, there is broad agreement that museums are informal and constructivist learning environments, whereby learning is freely directed by the learner, includes multiple entry points, and is highly personal to a visitor's complex identity (Falk & Dierking, 1992, 2000, 2016, 2018; Hein, 1998). Museums are seen as places where visitors feel and think, connect with objects and narratives, interpret collections through their own identities, and, among other things, learn (Baker, 2015; Hooper-Greenhill, 1992, 2000; Macdonald, 2006; Radywyl et al., 2015; Roschelle, 1995; Watson, 2015).

However, a key issue of such conceptualisations is that they struggle to support museum practice in understanding and catering for museum visitors. This is primarily because of the difficulty of evidencing constructivist museum visitor learning taking place (Falk & Dierking, 2016; Hein, 1998; Rennie & Johnstone, 2004). Thus, while constructivist understandings of visitor learning are useful theoretically, they struggle to provide practical support for museums in understanding visitor learning and engagement.

Visitor studies and evaluation in museums

Beyond these conceptualisations, extensive research into visitor behavior and identity has taken place in the field over the last 20 years. These studies explore the informal learning environment of museums and seek to capture the complexities of visitor identities (Falk, 2009; Hood, 1992; Pekarik et al., 1999; Rennie & Williams, 2002; Roberts, 1997; Schorch, 2015); the learning that is taking place in museum settings (Hooper-Greenhill, 2007); and visitor experiences within the museum (Alt & Shaw, 1984; Falk, 2009; Griggs, 1990; Hood, 1992; Hooper-Greenhill, 2007; Johnston, 1999; Lindauer, 2005; Miles & Tout, 1994; Pekarik et al., 1999; Rennie & Williams, 2002; Roberts, 1997; Schorch, 2015).

Similarly, exhibition evaluation practice has taken place within organizations as a means of documenting the relative successes and weaknesses of exhibitions. In practice, for both visitor studies and evaluation, the museum sector relies on a wide range of qualitative and quantitative methods to conduct front-end, formative, and summative research (Lindauer, 2005; Screven, in Bitgood & Loomis, 2012). Surveys, interviews, video-recordings, observations of visitor behavior and monitoring of exhibits are some of the most common methods (Nelson & Cohn, 2015) but are limited by the ingenuity of the individual researcher (Hein, 1998) and there is no single, fixed approach to evaluation (Diamond et al., 2016).

However, common evaluation practices are not without issues. There has been a trend for evaluations to produce overly positive results (Johanson & Glow, 2015) and to fail to produce real impact and learning for institutions (Davies & Heath, 2013). This is due to a lack of standardized methods, generalisable findings, the use of differing concepts of visitor profiles, and organizational constraints on making change. Therefore, while evaluation studies are relatively common, they are not necessarily impactful. Researchers have argued for the need to promote and engage practitioners in evaluation, because museum environments are increasingly

outcome-based and reflective environments with regard to understanding impact (Kubarek & Trainer, 2015). Practitioners are also challenged by a lack of time, resources, or experience, which means that evaluation is prioritized and embedded less than it ideally should be.

There are two key problems of evaluation and visitor studies that the MEUX Evaluation Method presented here seeks to address. First, it tackles the issue that little attention is paid to "what happens at the exhibit face" (Davies & Heath, 2013), meaning that evaluation and visitor studies gather a great deal of information about the complex visitor identity, but do little to explore how these visitors specifically interact with exhibits and exhibitions (Ardoin et al., 2016; Coelho et al., 2022; Heuken et al., 2021; Lucija et al., 2017; Ntamkarelou et al., 2017; Pecore et al., 2017; Phelan et al., 2020; Pietila, 2017; Robertson & Peterman, 2020; Shaby et al., 2019; Trucks et al., 2022). In recent years, there has been a shift toward more research studies that explore these specific interactions between museums and their visitors (e.g., Davis et al., 2023; Harrasser, 2015; Ji et al., 2023; Packer et al., 2022; Sneddon et al., 2021) but these often explore elements of the visitor experience (such as the use of smart phones or display labels) rather than a more comprehensive understanding. Thus, the MEUX Evaluation Method, based on the MEUX Model that conceptualizes museum–visitor interactions, seeks to address this problem by providing a means to capture information on the specific interaction between visitors and exhibits in a comprehensive way.

The second problem the MEUX Evaluation Method seeks to address is in relation to standardized and generalized evaluation results. Several authors have explored how a key limitation of evaluation reports is that they are solely focused on one particular exhibit or activity, utilizing individual methods and questions, which mean findings cannot be standardized or generalized into wider understandings of how visitors engage with exhibits (Fu et al., 2016; Nelson & Cohn, 2015; Peterman et al., 2020; Teasdale, 2022). Calls have been made for more standardizes approach (MacPherson et al., 2019; Ong & Ladenhead, 2015; Voiklis et al., 2023) and so the work presented here seeks to address such limitations.

These two problems, plus the limitation of constructivist theories not providing practical guidance to museums, are three key issues of current practice that the MEUX Model and Evaluation Methods seek to address. The Model acts as a clear conceptualisation of what happens "at the exhibit face," and this paper presents Evaluation Methods that allow museums to capture detailed information about these interactions. In developing these Evaluation Methods, the research explores available methods from the User Experience sector that may be adopted for use in museums.

User experience methods of evaluation

User Experience (UX) is concerned with the emotive and holistic experience a user has when interacting with a product, and Marc Hassenzahl's model of User Experience (2003) was integral to the development of the MEUX Model. User-Centred Design, a design philosophy that places the user at the center of product development, sits alongside UX principles. A crucial element of this is user research, which is undertaken to understand the complexity of potential users and provide concrete recommendations for product development (Van Kleef et al., 2005) using both qualitative and quantitative methods that capture and quantify user needs in a coherent way.

The toolbox of methods at the disposal of the user experience practitioner is diverse, ranging in detail, complexity, scope, and their qualitative and quantitative nature. Methods may have origins and uses outside the User Experience industry, but for the purposes of this research, focus has been concentrated on methods that are successfully utilized in User Research. However, not all are suitable or relevant to adapt for use in the museum sector to evaluate visitor experience with exhibits. This may be because they do not add value beyond existing

evaluation methods used in museum practice; because the method cannot be transferred to a museum context; because the method is based on the functional usability of a product rather than the more subjective user experience; or because the method involves specialist equipment and can be costly. Using these criteria, the research team identified a number of methods of UX research and evaluation that could be utilized for evaluating Museum Exhibit User Experience.

The first of these methods is the use of Semantic Differential scales, which has its origins in Osgood (1964), Hassenzahl et al. (2008), Karlsson and Wikstrom (1999), Laviea and Tractinsky (2004), Voss et al. (2003) and Wellings et al. (2008). Semantic Differentials are a method of measuring a concept or quality using a pair of opposing nouns, adjectives or phrases—such as 'good quality-bad quality,' 'hot-cold', or 'satisfied-unsatisfied.' Using a questionnaire, participants rate a product or service on a (usually) seven-point scale of the semantic differential item, providing a clear parameter with which to describe their perception. Ratings closer to the words at either end suggest more intense feelings, whereas ratings towards the middle of the scale indicate less intense feelings or judgment in relation to a product. This method is proposed for two key reasons. First, it is a method proposed by Hassenzahl, and since this research seeks to adopt Hassenzahl's model of User Experience, it seems natural to also utilize Hassenzahl's methods to capture information about product–user interactions. Second, this quantitative method is sophisticated in its ability to capture detailed information about the user–product interaction, but simple enough to be a self-reporting measure that does not require any particular expertise or resources in administering it. Thus, this is likely to improve the methods available in museum evaluation without requiring a significant increase in resources or staff skills.

Second is a group of methods that measure user emotions and affective responses to a product (Izard, 1979; Roseman, 1996; Scherer, 2005; Watson & Clark, 1994). The study and categorization of emotions has a long history within the psychology field, but the Geneva Emotions Wheel (Scherer, 2005) was selected by the researcher team as a well-established method containing 20 standard emotions as a means of assessing a particular element of the visitor experience. This was selected as it would specifically enable museums to capture information on the emotional elements of the visitor experience, which is an important element but one that is not often explored directly in evaluation.

Finally, there are several methods that seek to understand how and why certain qualities of a product related to user preferences and outcomes (Burmester et al., 2010; Liikkanen & Reavey, 2015; Zaman, 2008). These are more qualitatively based, and often rely on an interview technique called Laddering (Grunert & Bech-Larsen, 2005; Reynolds & Olson, 2001). Laddering is a process whereby, once a participant indicates they like or do not like something about a product, an interview asks a series of "why" questions to try and understand a user's reasoning. The method is called laddering because it seeks to increasingly extract reasoning on varying levels of abstraction. This "means-end approach" moves from product attributes to functional consequences (e.g., not feeling hungry) to psychological consequences (e.g., feeling good), to values or goals (Reynolds & Olson, 2001). The chain of reasoning produced by the laddering method means that researchers can understand how product attributes link to product qualities, link to experiential outcomes, beliefs, and values. This method is selected for two reasons. First, it is a means to capture the acknowledged complexities of the visitor identity in evaluation, and thus reflect the importance of the visitor in the museum–visitor interaction as extensively explored in the literature. Secondly, it enables complex, qualitative data, to be generalized and standardized by providing a means to lift specific visitor comments about exhibits into more generalized understandings of exhibits and visitor engagement.

These three methods: Semantic Differential scales, the Geneva Emotions Wheel, and interview Laddering, form the bases of the MEUX Evaluation Method. The following sections will discuss the development of the methods in a museum context; detail results from an evaluation

study of a temporary exhibition; explore results from a comparative study of three museum exhibits; and finally present a discussion exploring the benefits of the MEUX Evaluation Method.

METHODS

The purpose of conducting visitor evaluation underpinned by the concepts of the MEUX Model is to understand how visitors engage with a specific exhibit (King et al., 2023). In this paper, understanding visitor engagement with exhibits is conducted in two ways. The first is the in-depth exploration of the *Meat the Future* exhibition at OUMNH, exploring the MEUX evaluation data produced and how this can capture visitor impact. Second, there is a comparison study between *Meat the Future*, and two permanent exhibits, *Out of the Deep*, and *Passerines*, highlighting how the standardized MEUX evaluation methodology allows for rigorous comparison between exhibits of different types and an understanding of different visitor experiences.

Method development

The process of evaluation involved several methods (Table 1) to identify the highlight exhibition features; apparent exhibition character; whether any elements of visitor situation affected a visitor's perception of qualities; and final visitor outcomes.

The UX method that required the most development for this study was the use of Semantic Differential Word-Pair Scales. To develop a series of word-pairs relevant to museum exhibits, the research team conducted a method development and pilot study to produce a list of Semantic Differential Scales. The stages of this process were as follows:

1. Development of a first draft of word-pair Semantic Differential scales as informed by the MEUX Model and the data that informed its development (King et al., 2023).
2. Workshop with OUMNH staff to develop further Semantic Differential Scales based on ideal exhibit qualities.
3. Edit word-pairs to produce a second draft of Semantic Differential scales.
4. An online pilot study with a small number of museum visitors using Semantic Differential scales to evaluate two different presentations of exhibition material.
5. Statistical analysis of results assessing consistency of the word-pair scales.
6. Reduction of scales based on validity results, and further reduction to remove repeating or similar scales, to produce a final list of Semantic Differential scales.

In the process of reducing this initial list into the Pilot Study Draft, words were chosen to avoid the use of prefixes (e.g., "clear" and "unclear") in order to reduce positive visitor bias in evaluation (i.e., visitors reporting what they think the museum wants to hear). Therefore, there

TABLE 1 Summary of evaluation methods for MEUX model.

Section of MEUX model	Summary of evaluation methods
Exhibition Features	Laddering interviews and general comments from visitors
Exhibition Character: Pragmatic and Hedonic Qualities	Semantic Differential Word-Pair Scales Geneva Emotions Wheel
Visitor Situation	Visitor demographic data and motivations for visiting
Visitor Gains	Bad-Good Word-pair scale Likert scale agreements of outcome statements

TABLE 2 Pilot and final Semantic Differential Word-Pair Scales.

MEUX group	Word pair scales
Pragmatic Quality – Good Content	Confusing – Clear Rambling – Purposeful Messy – Organized Cluttered – Elegant
Pragmatic Quality – Good Interpretation	Boring – Interesting Irrelevant – Relevant Inaccessible – Accessible Distant – Engaging Dull – Inspiring Inclusive – Alienating
Both Pragmatic and Hedonic Quality – Coherence, Navigation and Physicality	Incoherent – Coherent Ambiguous to Follow – Intuitive to Follow Flat – Dynamic Rigid – Flexible Stuffy – Spacious
Hedonic Quality – Intellectual Engagement	Tedious – Stimulating Off-putting – Enticing
Hedonic Quality – Experience	Draining – Enriching[a] Condescending – Empowering[a] Forgettable – Memorable Removed – Relatable Isolating – Sociable Indifferent – Impactful Emotionally Removed – Emotionally Connected
Outcomes	Bad – Good (7-point Scale) 5-point Likert Scale from strongly disagree to strongly agree I feel I have understood the topic I feel I have engaged with the topic I feel I have learned something I have enjoyed myself I feel this has been a good use of my time I want to learn more about this topic I want to do something differently as a result of visiting this exhibition I want to return to the museum or another museum

[a]Scale more suited to Intellectual Engagement category in final study.

was a preference wherever possible to use two opposing but separate words to give each end of the Semantic Differential scale an identity in its own right so that a visitor assessment of an exhibit could be more objective and meaningful. The final list of the Semantic Differential scales are presented below (Table 2).

Evaluation study

The MEUX evaluation methods were used to evaluate three different exhibits at Oxford University Museum of Natural History. The full survey documents are provided as Appendix S1 to this paper. The survey was made up of the different evaluation methods identified above, covering visitor demographics, their motivation for visiting how they felt about the exhibit, how they perceived it, and how they rated various outcomes of their experience. Some of the word-pair scales were flipped to improve the validity of the survey but the directionality of scales was standardized prior to analysis to avoid confusion. Roughly 10% of visitors were

asked to take part in an extended version of the survey, where they had the ability to leave further comments in a short interview with the evaluator. This focused on highlighting particular exhibition features that stood out to them and influenced the ratings they had given.

Participants

The evaluation study was conducted at OUMNH in June 2021, following ethics approval from the University of Warwick Biomedical and Scientific Research Ethics Committee (reference 83/20-21) and the OUMNH ethics panel. The evaluator stood near the exhibit, and any visitors who had shown any form of engagement with the exhibit (e.g., beginning reading the exhibit) were asked to take part in the evaluation study when they had completed their engagement. This method was chosen, rather than random sampling, because the number of visitors engaging with each display was small enough to be able to include every visitor. If visitors were in a group, the whole group were invited to take part, and participants were self-selecting. Due to ethical considerations, only adults were able to take part, but there were many instances where parents completed the survey on behalf of their children. Having been asked to take part, visitors completed a small consent form and were then provided with full instructions to fill out their survey. The principal researcher, with a team of museum volunteers, evaluated three different museum exhibits over a period of 3 weeks. The exhibits were as follows (Figure 2):

1. *Meat the Future*: a temporary exhibition on display between May 2021 and June 2022. The exhibition presented the latest scientific research on the environmental and health impacts of meat and dairy consumption (online exhibition). Due to COVID-19 restrictions, the exhibit was designed to enable social distancing, with minimal touchable interactives.
2. *Out of the Deep*: a large, modern and permanent exhibit displaying the skeletons of two Jurassic plesiosaurs, which are large fossil marine reptiles (online display).
3. *Passerines*: a traditional, taxonomically structured, 'glass case' display of passerine (perching) birds created in the 1990s.

Data collected were interpreted using various quantitative and qualitative methods, using IBM SPSS Statistics and QSR NVivo, respectively. The visitor demographic data and selected emotions were analyzed for simple frequencies, and the Semantic Differential scales were subject to mean calculations. Further analysis was then performed by calculating cross-tabulation results between demographic data and semantic differential scales, to assess whether visitor identity had any effect on their ratings. Correlations were performed between the Semantic Differential scales to assess the relationships between the scales and outcomes for visitors. Qualitative data were thematically coded to identify common themes and features highlighted by visitors. This analysis then created a MEUX summary of each exhibition, bringing together the evaluation data to provide a snapshot perception of the exhibition from the visitor perspective.

These summaries allowed for comparison between different exhibits, exploring how each visitor experience is constructed in different ways for different types of exhibits. A more direct comparison was made through statistical analysis of the various mean scores of Semantic Differential scales and Outcome statements of different exhibits. The results were compared using Kruskal–Wallis one-way analysis of variance (due to the nonparametric nature of the data) to identify statistical differences between the exhibits. A significant result ($p < 0.05$) demonstrated a significant difference in ratings between the three exhibits. This range of analytical techniques allowed for the building of a picture of the mechanisms of visitor experience for each exhibit, using evidence and data to create holistic understanding, rather than it being presented in its own right. Below are the results of the in-depth study of *Meat the Future*, and a comparative study of three exhibits.

FIGURE 2 Exhibit images (see Images document). (a) Meat the Future. (b) Out of the Deep. (c) Passerines. [Colour figure can be viewed at wileyonlinelibrary.com]

RESULTS

Case study: Meat the Future

The evaluation of *Meat the Future* involved the collection of 242 individual surveys and 22 short interviews. The results explored below build a picture of the *Meat the Future* exhibition as understood and perceived by visitors, using quantitative and qualitative data, which is summarized within the MEUX Model framework.

Demographic data—What types of visitors were there?

The first section of the evaluation survey was to gather information about visitors of *Meat the Future*. This included gender, age group, who they were visiting the exhibition with, how often they visited museums generally within a year (Figure 3), and why they had visited the museum that day (Figure 4). Frequencies for each option were counted and are presented below. This information provides an initial understanding of the types of visitors who were engaging with

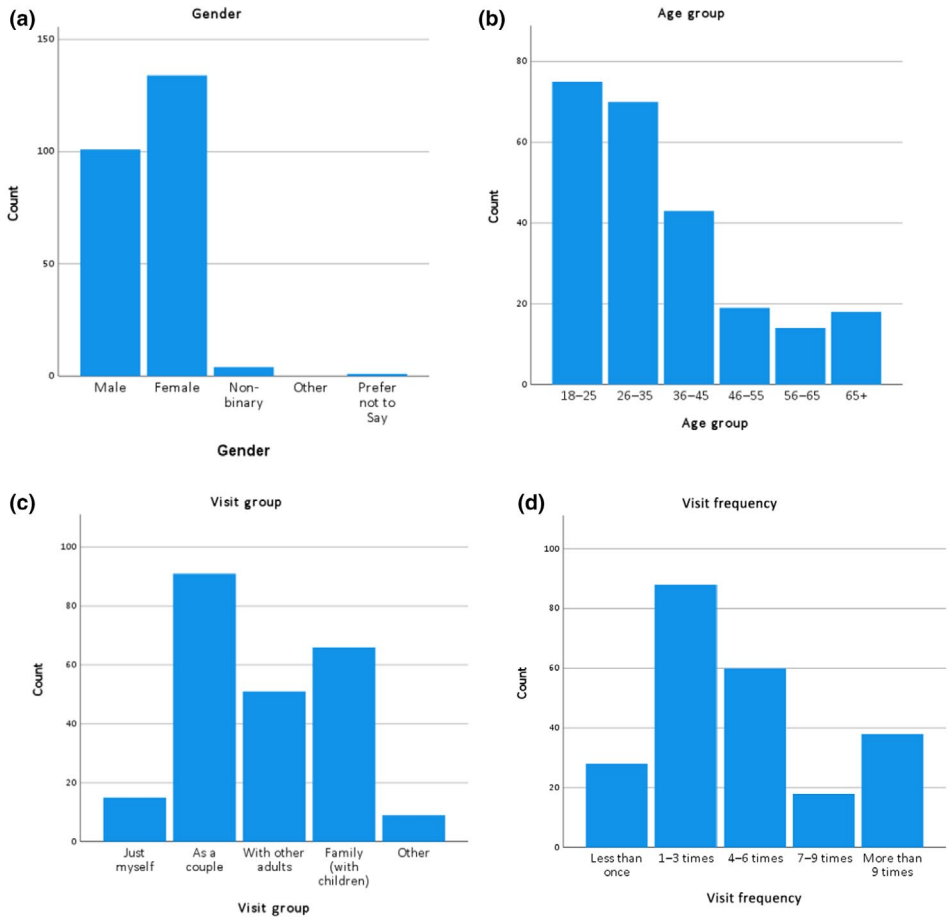


FIGURE 3 Demographics of visitors to Meat the Future. (a) Gender. Visitors were asked 'What is your gender?' (b) Age Group. Visitors were asked 'What is your age group?' (c) Visit Group. Visitors were asked 'Who are you visiting the museum with today?' (d) Visit Frequency. Visitors were asked 'How often do you visit museums in general in a normal year?' [Colour figure can be viewed at wileyonlinelibrary.com]

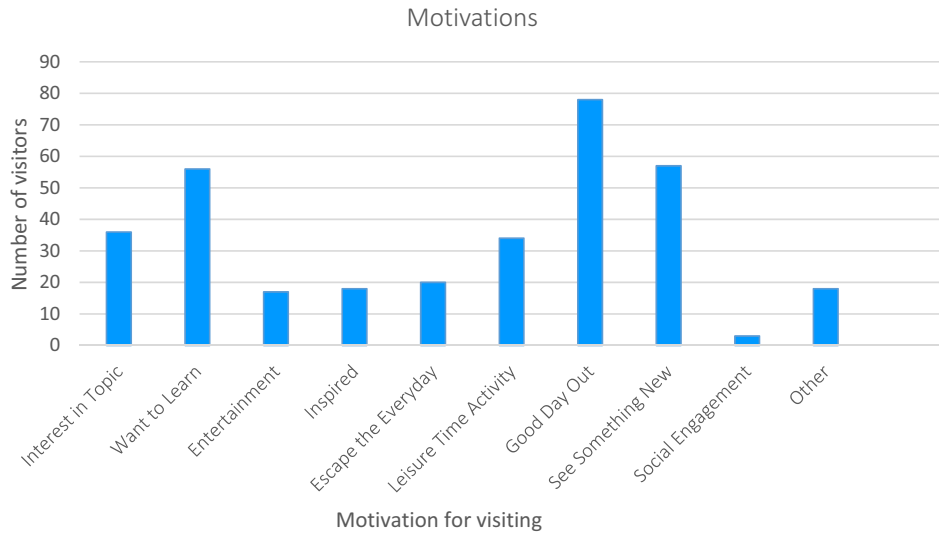


FIGURE 4 Motivations. Visitors were asked 'What was the strongest motivation for visiting the museum today?' [Colour figure can be viewed at wileyonlinelibrary.com]

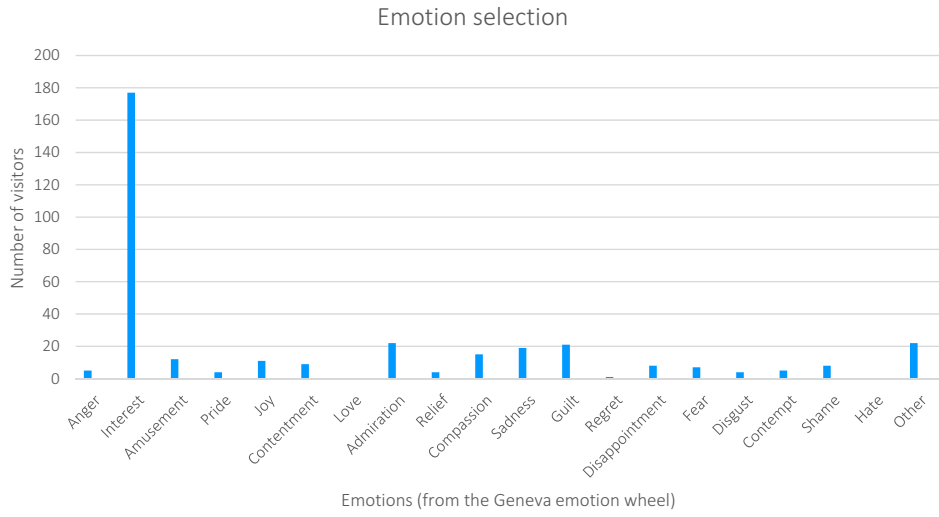


FIGURE 5 Emotion selections. Visitors were asked 'Which emotion do you associate most strongly with when considering your visit to this exhibit?' [Colour figure can be viewed at wileyonlinelibrary.com]

Meat the Future but was also useful for further analysis of how visitor identity affects perceptions of the exhibit (see cross-tabulations, below).

Emotion selection—How did people feel about the exhibition?

Frequencies were counted for emotion selections that visitors made from the Geneva Emotions wheel. Results (Figure 5) show that *Meat the Future* made visitors feel quite negative and have strong emotions, such as compassion, sadness, and guilt. One visitor said "I feel slightly guilty but I don't want to acknowledge that – I have to live with it every day."

However, there were also some positive emotions, such as admiration, pride, and joy. One visitor commented they felt "admiration for a brilliantly put together, very informative exhibition. My daughter has just gone vegan and as a result of this exhibit we have decided to cut beef from our diet for the rest of the family." These results provide preliminary understanding of how visitors perceived the exhibit but can be used in further evaluation as discussed below (see cross-tabulations).

Semantic differential scales results—How do visitors perceive the exhibition?

In analysis of the Semantic Differential scales visitors used to rate the exhibition, ratings were translated into scores from 1 to 7, with 1 indicating the less positive word, and 7 indicating the more positive word (Figure 6). A mean score of results was then calculated, with 4.0 being a neutral rating. In analysis of these results (Figure 7), each mean score is relatively high and leans toward the positive term, but a more useful approach is to assess the scores in relation to each other. We can determine which qualities of the exhibition were most important and stood out the most to visitors by comparing scales within a group.

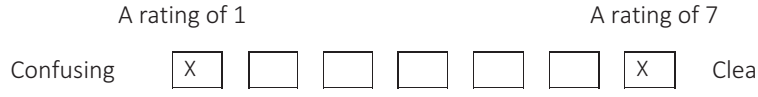


FIGURE 6 Example visitor rating using semantic differential word pair 7-point scale.

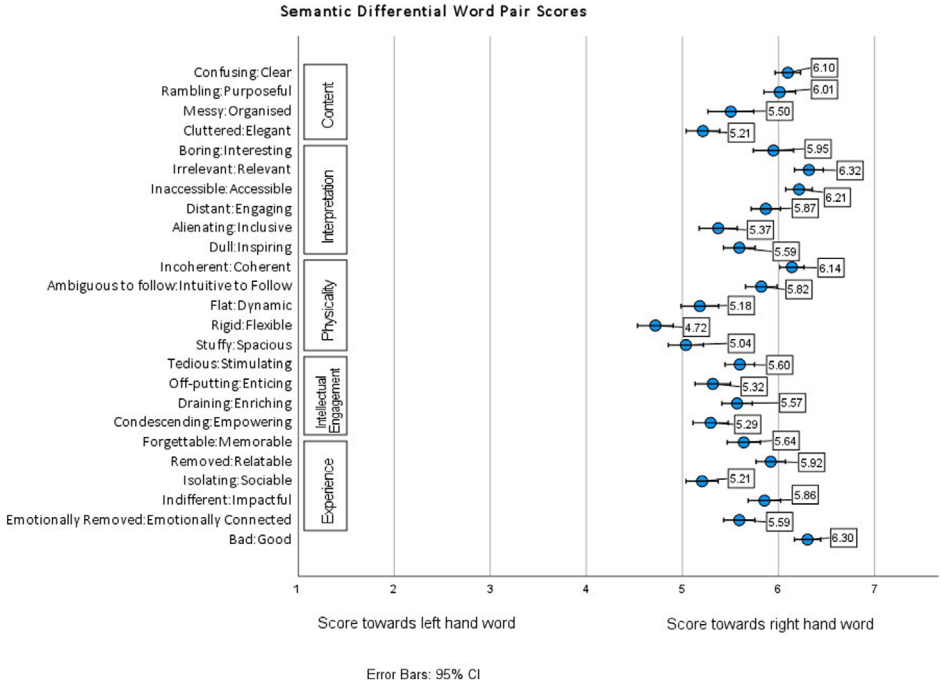


FIGURE 7 Mean results of Semantic Differential Word-Pair scores with a 95% Confidence Interval. Despite all results being very positive and towards the right-hand word of the pair, by assessing the relative differences between results we can build a picture of the key qualities of the exhibit being evaluated. [Colour figure can be viewed at wileyonlinelibrary.com]

By identifying the scales that performed the strongest for each of the "content," "interpretation," and "experience" groups, visitors perceived *Meat the Future* to be about an important and relevant (6.32) topic and that the exhibition was driven by its intellectual, rather than experiential nature. The exhibition was considered clear (6.10) and purposeful (6.01); interesting (5.95), relevant and accessible (6.21); and relatable (5.92) and impactful (5.86) to visitors. One visitor provided a comment that "the exhibit related to something we do every day so was more relevant and thought-provoking," which summarizes the perception visitors got from *Meat the Future*.

Considering the results overall, there were slightly lower scores in the "navigation and physicality," "intellectual engagement," and "experience" groupings, again expressing that the focus of the exhibition was on its topic and visitor engagement with this information. The low physicality results, especially on the "rigid-flexible" and "flat-dynamic" scales, was most likely due to COVID-19 restrictions which meant the exhibition was designed to allow for social distancing within the long, narrow gallery it was presented in. A lack of varied learning styles contributed to the slightly lower "intellectual engagement" scores, and it is interesting that the two highest experience scores, as "relatable" and "impactful" speaks directly to the nature of the topic of *Meat the Future*: meat consumption and individual diets. Visitors perceived the exhibition to be very information led rather than experience led, which contrasts with how visitors usually encounter information concerning meat consumption. One visitor commented that "the exhibition is more information based; it is less sensationalist than TV programmes," with another stating that "we've spoken about this subject recently so we've moved on from the 'terror' of the situation."

Overall, the semantic differential scales identify that visitors perceived *Meat the Future* as a topic-driven exhibition, providing information on a subject that is relevant and important, but providing less of an experience or varying engagement for them.

Outcome results—What do visitors get out of the exhibition?

Visitors rated their level of agreement with eight statements about what they got out of the exhibition on a 5-point Likert Scale, ranging from "strongly disagree" (score of 1) to "strongly agree" (score of 5). Again, mean scores were calculated, and the results are presented below (Figure 8).

In a similar way to the Semantic Differential scales, assessment was undertaken not of the scores in their own right, but in comparison with each other. In these results, there are some small differences between the understanding (4.44) and learning (4.35) outcomes, which are rated higher, and the engagement (4.24) and enjoyment (4.25) outcomes, which are rated slightly lower. Again, this demonstrates that the focus of *Meat the Future* was on providing information and learning, rather than experience, for visitors, and enabling them to make their own evidence-led decisions. The slightly lower score for enjoyment corroborates the negative emotion selections discussed above and is understandable considering the nature of the exhibition topic. The ratings for wanting to learn more (4.10) and do something differently (3.69) are fairly high, suggesting that the information has made an impact on visitors and they were keen to engage with the topic further. Visitors who were interviewed provided detail for this desire to change their behavior, and this ranged from a statement that they wanted to ($n=4$), that they would consider it ($n=3$), or that they would do more research ($n=2$). One visitor said "I find it inspiring but I'm already vegan so not compelled to change," whereas another commented that "I will do more meal prep and more meat free days. I'll go to the butchers rather than the supermarket."

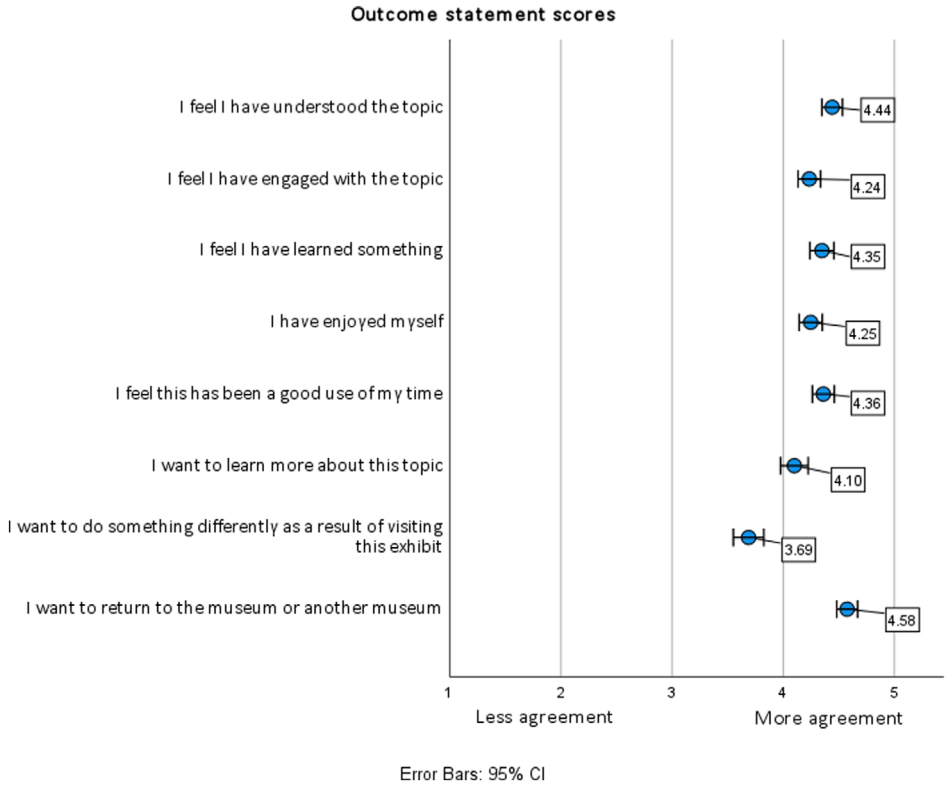


FIGURE 8 Outcome statement scores. A higher score indicates more agreement with the statement. Despite all results receiving high mean scores, by assessing the relative difference between outcomes we can build a picture of what visitors gained from the exhibit being evaluated. [Colour figure can be viewed at wileyonlinelibrary.com]

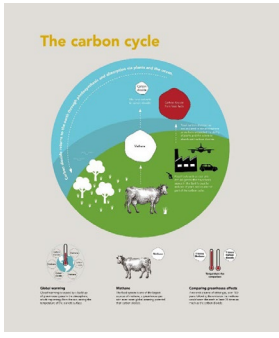
Interview data—What exhibition features are important?

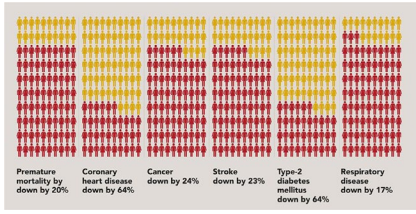
Visitors who took part in the interview section of the evaluation ($n = 22$) were asked which elements of *Meat the Future* stood out to them the most, and which elements influenced the ratings visitors gave (see Figure 9 for images). The data were thematically coded, with a new code added for each new element a visitor mentioned, and the table below (Table 3) shows the most popular exhibition elements based on the number of visitors who mentioned them.

Results show that visitors were drawn to the statistical evidence and scientific information that were included in the exhibition, which align with the Semantic Differential and outcome results discussed above. These features also demonstrate the importance of how the information is presented, with graphs, charts, and visual data presentations being among the most commented elements of the exhibition. The interview data from visitors also provided understanding of various themes concerning the exhibition, which emerged from coding the interview comments into themes. Visitors felt that the exhibition was a personal topic and something that they were conscious of prior to their visit, and that it was empowering and informative. A sense of choice and empowerment was clearly important for visitors as it is such a personal topic, with one visitor commenting that "I was already trying [to eat less meat] but now with specific subject knowledge I feel empowered to change."

(a) Meat consumption Chart

(b) Visuals and graphics



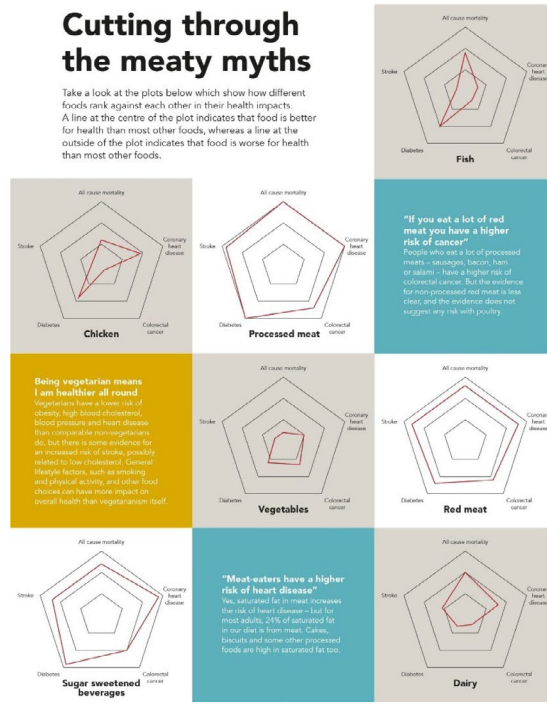


(c) Food images and examples

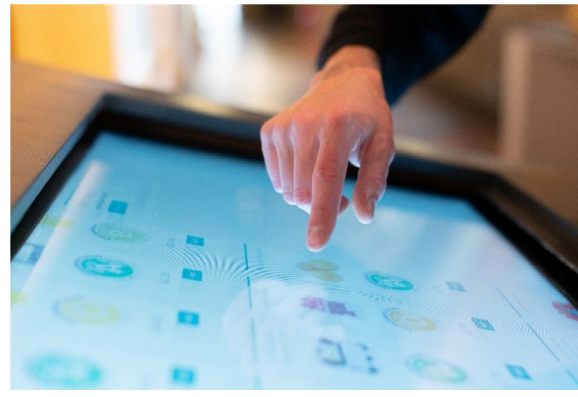
FIGURE 9 Meat the Future Exhibit Features accompanying images. [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 9 (Continued)

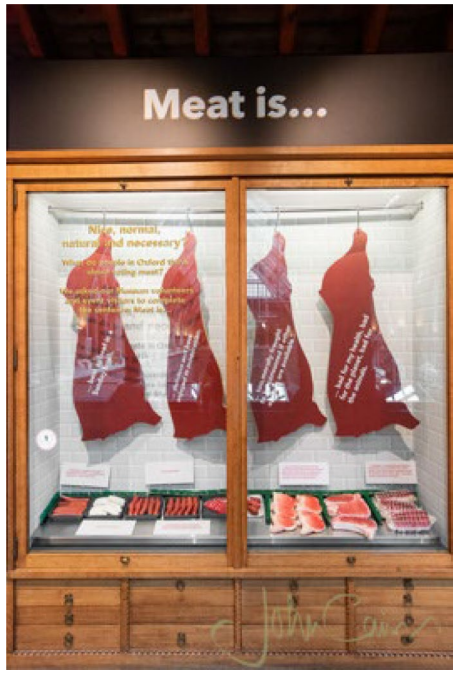
(d) Health impact diagrams



(e) Environmental labels and shopping task



(f) Meat Is



(g) Emissions Chart

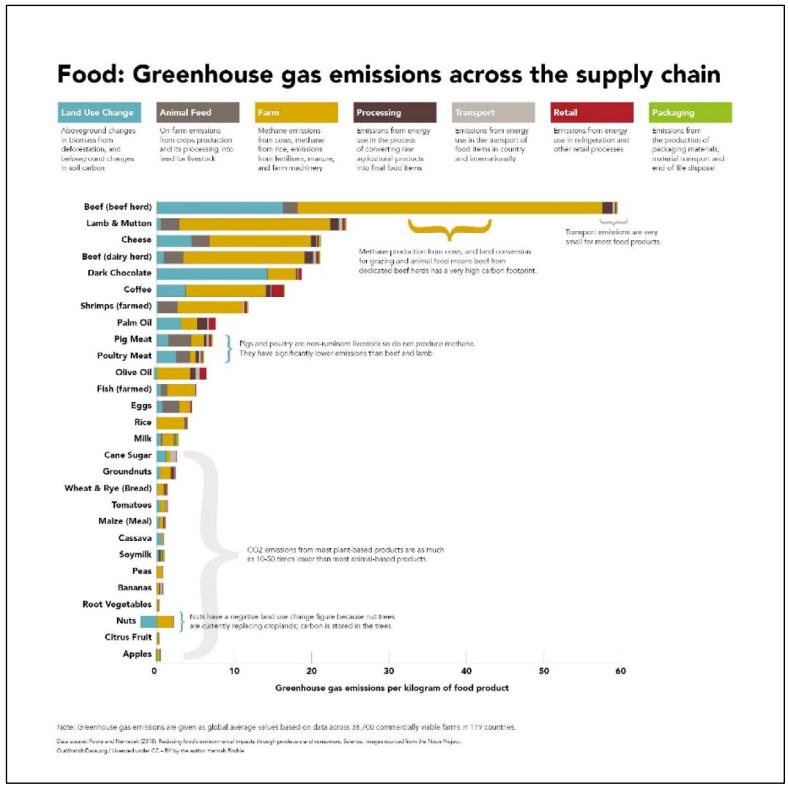


FIGURE 9 (Continued)

TABLE 3 Most popular features of the *Meat the Future* exhibition.

Exhibition feature and description (with accompanying images)	Frequency
Meat Consumption Chart (A) *Visual chart using stacks of burgers to represent meat consumption in various countries*	11
Visuals and Graphics (B) *The various infographics, charts, and visual elements presenting information*	9
The scientific evidence concerning meat consumption presented *The scientific information featured in the exhibit*	8
General environmental impact *Information concerning the environmental impact of meat*	5
Food images and examples (C) *Examples of low-meat and plant-based meals*	4
Health impact diagrams (D) *Series of diagrams showing the health impact of various foods*	4
Environmental labels and shopping task (E) *Packaging label displaying environmental impact of food and a digital interactive where visitors can conduct an online shop and see the environmental impact of their choices*	3
'Meat Is…' display (F) *Display of various opinions about what 'meat is…'*	3
Emissions Chart (G) *Chart showing emissions data of various food types*	2

How could the exhibition be improved?

Naturally, visitors also provided more negative comments and offered suggestions of improvement for *Meat the Future*. Some of these were more technical and functional changes, such as the text needing to be bigger ($n=2$) and that some clarifications were needed ($n=1$). Others related to the interactivity of the exhibition, with some visitors commented that they generally wanted more interactive elements ($n=2$) and others wanting more variety for different learners, such as more audio ($n=1$) and visual ($n=1$) elements. These comments are in line with the Semantic Differential results, where the results for the intellectual engagement and experience groups are slightly lower. One visitor highlighted the linear and flat exhibition space (a by-product of COVID-19 social distancing) being difficult for a visual learner, and another wanted generally more information to raise awareness of the issue. Some visitors wanted more plant-based alternatives offered ($n=2$) while others were concerned with the health impacts of these alternatives and would have liked to see nutritional information on them ($n=3$).

Cross-tabulations—Does visitor identity affect perceptions or outcomes?

The results so far have identified the features, qualities, and visitor outcomes for the *Meat the Future* exhibition. Now, in conducting further statistical calculations, we can determine more detailed understanding of the visitor experience, and capture trends from a wide range of heterogeneous visitors. The first of these is performing cross-tabulation calculations to assess whether different types of visitors perceived and rated the *Meat the Future* exhibition differently. Contingency tables were used to assess whether there were differences in results for each different type of visitor demographic, taking sample size into account. Results were calculated for Gender using Fisher's exact test, and for Age Group, Visit Group, Visit Frequency, Motivation, and Emotional Selection using the Monte Carlo method based on 10,000 samples

TABLE 4 Significant cross-tabulation results between visitor demographics and semantic differential scale and outcome scale ratings.

Visitor demographic	Scale or outcome	Test result ($p < 0.050$)
Gender	Confusing-Clear	0.039
Gender	Rambling-Purposeful	0.020
Gender	Messy-Organized	0.030
Gender	Cluttered-Elegant	0.038
Gender	Irrelevant-Relevant	0.017
Gender	Distant-Engaging	0.043
Gender	Dull-Inspiring	0.000
Gender	Ambiguous to Follow-Intuitive to Follow	0.027
Gender	Flat-Dynamic	0.001
Gender	Tedious-Stimulating	0.038
Gender	Off-putting-Enticing	0.000
Gender	Draining-Enriching	0.012
Gender	Condescending-Empowering	0.043
Gender	Removed-Relatable	0.031
Gender	Emotionally Removed-Emotionally Connected	0.019
Gender	Understood the Topic	0.019
Gender	Enjoyed Myself	0.031
Gender	Want to Learn More	0.014
Gender	Want to Do Something Different	0.001
Gender	Want to Return	0.002
Visit group	Dull-inspiring	0.003
Visit group	Ambiguous to follow-Intuitive to follow	0.026
Visit group	Draining-Enriching	0.016
Visit group	Isolating-Sociable	0.009
Visit frequency	Confusing-Clear	0.023
Visit frequency	Inaccessible-Accessible	0.001
Visit frequency	Flat-Dynamic	0.011
Visit frequency	Understanding of topic	0.001
Visit frequency	Want to return	0.001
Motivation for visiting	Engagement with topic	0.050
Emotion selection	Distant-Engaging	0.011
Emotion selection	Dull-Inspiring	0.048
Emotion selection	Flat-Dynamic	0.002
Emotion selection	Stuffy-Spacious	0.016
Emotion selection	Off-putting-Enticing	0.001
Emotion selection	Draining-Enriching	0.005
Emotion selection	Removed-Relatable	0.015
Emotion selection	Isolating-Sociable	0.002
Emotion selection	Bad-Good	0.044
Emotion selection	Engaged with Topic	0.000

TABLE 4 (Continued)

Visitor demographic	Scale or outcome	Test result ($p < 0.050$)
Emotion selection	Learned Something	0.001
Emotion selection	Enjoyed	0.012
Emotion selection	Good Use of Time	0.012
Emotion selection	Different Behavior	0.015

Note: These results show how different types of visitors engaged with the exhibit differently.

to a 99% Confidence Interval. Significant results of $p < 0.050$ are identified and analyzed for wider trends (Table 4).

In analysis of results by visitor gender, there were 15 scales and five outcomes that produced significant results, and in all except one scale, the results showed that females rated the exhibition more highly than their male counterparts. It is uncertain why these results emerge, but the visitor quote below may provide a male attitude toward meat-eating:

> I'm a meat eater, but it feels like you're being told to do something different, there's that pressure. But it's informative and stuff I knew already. I understood the damage. But it helps form an opinion rather than jumping on the bandwagon.

For visitors of different ages, there were no significant results reported, suggesting this factor did not affect visitor perceptions. However, although *Meat the Future* was targeted at visitors over the age of 16, there were some significant results when considering Visit Group, which included families with children. Family groups found the exhibition less intuitive to follow and less enriching than other groups, highlighting their lack of access and engagement with an exhibition not strictly designed for them. Nevertheless, it is also interesting to note that family groups rated the exhibit as both inspiring and sociable more than other groups. This suggests that while not being able to engage with the exhibition on a more detailed level, there was general engagement with the topic and an ability for family visitors and children to find inspiration in reducing meat consumption, and ability to discuss this sociably as a family.

For Visit Frequency, there were three scales and two outcomes that produced statistically significant results. The results demonstrate that the exhibition was engaging for people who visited museums less frequently as well as more frequently, as less frequent visitors rated the exhibit as clearer, more accessible, and more dynamic. There was also a significant result for the statement "I want to return to the museum," with more frequent visitors agreeing more strongly with this, but this is unlikely to be due to the exhibition itself and rather the nature of the visitor.

In analyzing results concerning a visitor's motivations for visiting the museum, the outcome of engagement ("I feel I have engaged with the topic") produced significant results. Visitors who were visiting for intellectual and social reasons had more engagement, but visitors who were there to escape the everyday had less engagement. This may reflect how the *Meat the Future* topic is very current, timely, and relevant, and its close association with everyday life did not chime with visitors' wishes to escape from this.

For five of the outcome statements, results show that a visitor's emotion selection had an impact on their ratings. As mentioned above, the key emotions selected in response to *Meat the Future* were "interest" and the group of somewhat negative emotions consisting of "sadness," "guilt," "shame," and "fear." There is a clear trend in the results. 'Interested' visitors rated the outcomes of enjoyment, and the exhibition being a "good use of my time," more highly.

However, visitors who said they felt guilty, sad, or shameful rated their engagement and learning more highly. This may reflect how visitors who had more emotional engagement with the exhibition felt negative emotions such as guilt or sadness. While this represents meaningful engagement with exhibition content, it cannot be classified as enjoyable for the visitor. By contrast, those with less emotional engagement found *Meat the Future* a more enjoyable experience. Most interestingly, visitors who felt 'guilty' were more likely to want to do something differently as a result of visiting. This is extremely important for an exhibition like *Meat the Future*, as the emotional engagement of visitors has led to a statement of intent that they will change their behavior on meat consumption.

Overall, performing cross-tabulation analysis enables us to understand how different types of visitors interact with *Meat the Future*, thus building a more nuanced and detailed picture of visitor experience and providing trends of perceptions from different visitor groups.

Correlations—What qualities influence outcomes the most?

The second statistical analysis performed on the evaluation data was to examine correlations between Semantic Differential scales and Outcomes. This means we can identify which exhibit qualities have the most influence on visitor outcomes and begin to understand the mechanisms of experience for visitors of *Meat the Future*. Due to the non-parametric nature of the data, as identified by the Shapiro–Wilk test of normality, bivariate correlational analysis was performed and a correlation coefficient was calculated using Spearman's Rho. Generally, a correlation below ±0.3 is weak, between ±0.3 and 0.7 is medium, and above ±0.7 is strong (Asuero et al., 2007). Correlations between Scales and Outcomes were calculated, and each outcome presented with the three most strongly correlated scales (Table 5).

As the results are all positive correlations, this means that the stronger agreement a visitor had with an outcome statement, the more likely they were to select a rating toward the positive end of the word-pair scale. For example, the more clearly a visitor rated the exhibition, the more they agreed that they had understood the topic. The correlational analysis helps us to understand what features and qualities of *Meat the Future* contribute to visitors' overall judgment and a variety of outcomes. Interestingly, different scales correlate with different outcomes, showing that various factors are at play in the exhibition at any one time, contributing to different outcomes. For example, visitors judged the exhibit as "good" when they felt it was stimulating, inspiring and impactful, with similar experiential scales correlating with the outcomes of enjoyment. Alternatively, understanding the topic correlated with the exhibit being "clear," "intuitive to follow," and "relevant," highlighting how the more objective scales concerned with content and presentation are important for understanding. Engagement came from the exhibition being "impactful," "relevant," and "engaging," which were some of the highest-rated scales and perceptions for *Meat the Future*. Perhaps most interestingly, wanting to do something differently as a result of the exhibition stems from qualities such as "impactful," "empowering," and "inspiring."

In these results, the majority of correlation coefficients reported are categorized as medium strength. This means that while there is a relationship between scales and outcomes, they are not as clear cut or direct as they could be, and it is likely that there are other factors at play. This may be unsurprising for understanding museum visitor experience, as there are various *Affecting Factors* at play that mediate a visitor's perception of an exhibition and subsequent outcomes, as conceptualized in the MEUX Model. It is not possible to capture all of these factors and so it is natural for relationships between qualities and outcomes to be less homogenous and strong in statistical results.

TABLE 5 Correlation results between each outcome statement and the three strongest correlating word-pair scales.

Scale	Correlation coefficient
Bad-good	
Tedious-Stimulating	0.560
Dull-Inspiring	0.548
Indifferent-Impactful	0.528
I feel I have understood the topic	
Confusing-Clear	0.414
Ambiguous to Follow-Intuitive to Follow	0.401
Irrelevant-Relevant	0.378
I feel I have engaged with the topic	
Indifferent-Impactful	0.499
Irrelevant-Relevant	0.489
Distant-Engaging	0.463
I feel I have learned something	
Indifferent-Impactful	0.459
Flat-Dynamic	0.414
Dull-Inspiring	0.409
I have enjoyed myself	
Draining-Enriching	0.455
Tedious-Stimulating	0.442
Dull-Inspiring	0.435
I feel this has been a good use of my time	
Dull-Inspiring	0.507
Tedious-Stimulating	0.488
Draining-Enriching	0.451
I want to learn more about this topic	
Irrelevant-Relevant	0.465
Tedious-Stimulating	0.443
Dull-Inspiring	0.403
I want to do something differently as a result of visiting this exhibition	
Indifferent-Impactful	0.456
Condescending-Empowering	0.431
Dull-Inspiring	0.387
I want to return to the museum or another museum	
Dull-Inspiring	0.242
Forgettable-Memorable	0.232
Distant-Engaging	0.216

Note: All reported results are significant. As the correlations are all positive, this means that the outcome correlates with the right-hand term in the pair. For example, the exhibit being rated as 'good' correlates with it also being rated as stimulating, inspiring, and impactful.

Comparative study: Three exhibits

The MEUX evaluation methods use quantitative and qualitative data to provide an overall insight into the perceptions that visitors have of the exhibit. One of the key benefits of this method, however, is that the standardized methodologies also allow us to directly compare exhibits and assess the different impacts that exhibits have on visitor experience. This second study therefore compares the results from three different exhibits at OUMNH: *Meat the Future*, which we have already discussed, plus *Out of the Deep* and *Passerines*, two permanent yet different displays at the museum. *Out of the Deep* evaluation collected 237 surveys and 23 interviews, while *Passerines* evaluation collected 235 surveys and 22 interviews. The quantitative and qualitative data collected for each exhibit, and explored for *Meat the Future* above, is summarized in an exhibit-specific MEUX framework, as detailed below (Tables 6–8).

By comparing these three exhibit summaries, it is clear how the three exhibits are constructed in very different ways and have very different visitor experiences. *Meat the Future* was intellectually driven concerning a societally important topic of current importance, whereas *Out of the Deep* is more experientially driven with strong interactive elements, and *Passerines* is a 25-year-old, traditional taxonomically themed museum case (now dismantled as part of a redisplay programme) that was reliant on prior subject knowledge and interest for visitor engagement as it does not provide opportunities for connection or interaction in its own right.

The use of the MEUX Model and evaluation methods allow for clear and concise comparison between these three contrasting exhibits. Statistical Analysis of Variance was performed in order to identify statistical differences between the three exhibits. Below are the results reported for the "Interpretation" group of Semantic Differential scales (Figure 10 and Table 9), and all outcome statements (Figure 11 and Table 10).

The results provide us with the ability to make a direct comparison between the exhibits. For example, the scale "irrelevant-relevant" has produced a significant result, and in comparing the mean scores on the graph we can see that *Meat the Future* is statistically considered to be the most relevant of all the exhibits evaluated. In contrast, for the "alienating-inclusive" scale, we can see that *Out of the Deep* is considered statistically more inclusive than the other two exhibits. Instead of just perception or opinion about the differences between exhibits, these results provide us with solid statistical evidence. For the outcome statement results, we can see that *Out of the Deep* was the most enjoyable exhibit, but *Meat the Future* gave people a desire to do something differently. Combining these data with the information gathered about each exhibit, it is possible to understand the reasons why a particular exhibit has scored better on a scale. For example, we know that visitors recognized how personal the topic of meat consumption was, and so we know this has contributed to its sense of relevance. For *Out of the Deep*, many visitors appreciated the small portholes in the exhibit that children could look through at their own height, which has contributed to the sense that the exhibit is inclusive.

These results demonstrate how the MEUX Model and evaluation methods produce generalisable, standardized, and comparable data of three very different museum exhibits. It provides nuanced and detailed understanding of how different exhibits and visitor experiences are constructed, and how these compare to each other. It does not necessarily mean that one particular exhibit is better than another overall, but rather demonstrates how all three exhibits engage visitors in very different ways. One exhibit may perform better on a particular scale or outcome, but this scale may be less relevant for another exhibit. Essentially, the MEUX methodology allows museum practitioners to capture, from visitors, the extent to which an exhibit is doing what it was intended. If *Meat the Future*'s aim was to inform and empower people to make decisions about their diets, the MEUX method captures how well this has been achieved. For the *Out of the Deep* or *Passerines* exhibits, this decision making is less relevant, and the data reflect this. It does not mean it has failed, just that this measure was less relevant.

TABLE 6 MEUX summary of *Meat the Future*.

Exhibition features	Exhibition character	Visitor situation	Outcomes
Content Facts and science presented Meat consumption chart Environmental shopping *Presentation* Strong visuals *Functionality* Slightly flat space Text too small *Interaction* Environmental shopping task Health impact diagrams Needs wider learning styles	*Pragmatic Qualities* Balanced view presented Relevant and topical issue *Hedonic Qualities* Can induce negative emotions Personal, makes you consider your own impact	*Affecting Factors* Current diet of visitor Strong prior knowledge and interest Aimed at adults, not children *Motivations* Interest in topic – but unexpected to see in museum A good day out	*Visitors to Know* More information about impact of meat consumption *Visitors to Feel* Some strong negative emotions, but not necessarily enjoyment *Visitors to Do* Eat less meat Some visitors felt they were already doing lots

TABLE 7 MEUX summary of *Out of the Deep*.

Out of the deep MEUX summary			
Exhibition features	**Exhibition character**	**Visitor situation**	**Outcomes**
Content Skeleton remains presented *Presentation* Underwater theme *Functionality* Size of skeletons portrayed *Interaction* Peep holes for children to look through, makes them feel underwater	*Pragmatic Qualities* Clear and interesting topic *Hedonic Qualities* Sense of connection, realism, experience of being in the water with creatures	*Affecting Factors* Being with children, quick reading of information Not isolated from rest of exhibits, is part of the museum *Motivations* Intellectual engagement Bringing children to learn	*Visitors to Know* Good understanding of topic *Visitors to Feel* A sense of wonder, fascination, and awe *Visitors to Do* No direct actions related but a sense of contentment and satisfaction

TABLE 8 MEUX summary of *Passerines* display.

Passerines MEUX summary			
Exhibition features	**Exhibition character**	**Visitor situation**	**Outcomes**
Content Colorful birds, nests *Presentation* Birds in flight Confusion about titles and descriptions *Functionality* Text-based *Interaction* Very little	*Pragmatic Qualities* Only partly clear and interesting *Hedonic Qualities* Very little experience for visitors	*Affecting Factors* Prior interest in topic matters a great deal *Motivations* Intellectual motivations very important	*Visitors to Know* Basic information about birds but some confusion about 'Passerines' label *Visitors to Feel* Only some emotions from visitors with prior interest *Visitors to Do* Visitors keen for direct action relating to environment

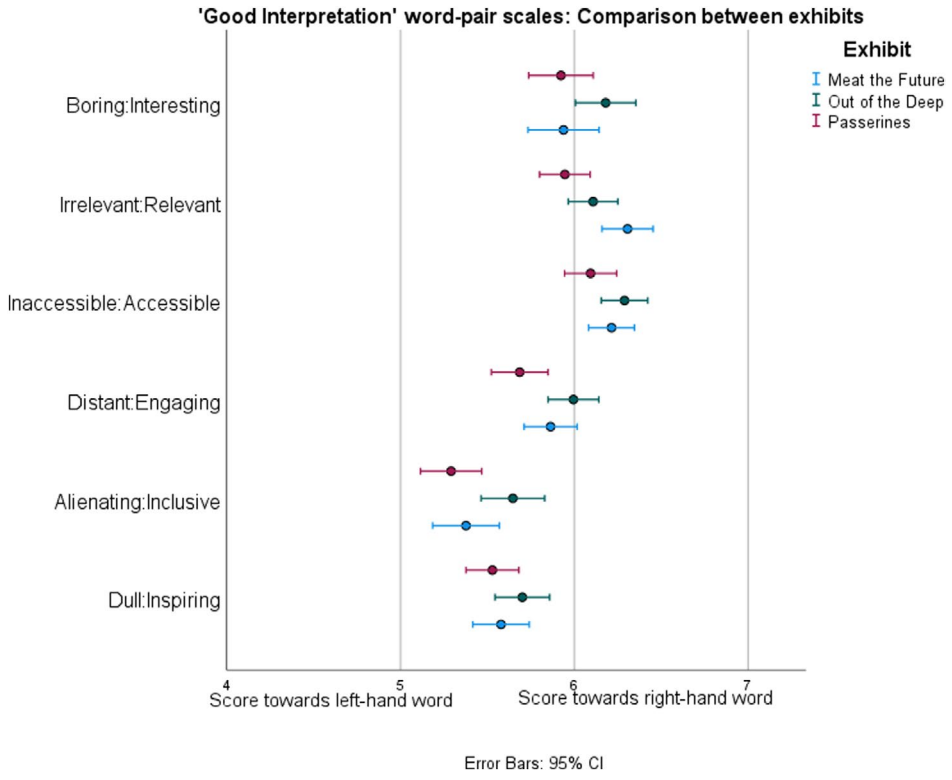


FIGURE 10 Comparison of 'Interpretation' Word-Pair Scale results for three evaluated exhibits. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 9 Statistical results for Kruskal-Wallis one way analysis of variance of Interpretation word-pair scales between three different exhibits.

Scale	Kruskal-Wallis test statistic	Difference
Boring-Interesting	0.044	Different
Irrelevant-Relevant	0.000	Different
Inaccessible-Accessible	0.137	Not different
Distant-Engaging	0.035	Different
Alienating-Inclusive	0.004	Different
Dull-Inspiring	0.245	Not different

DISCUSSION

Both the case study and comparative study results presented above demonstrate the multiple benefits of utilizing the MEUX Evaluation Method to conduct visitor evaluation as described. Evaluation is rigorous, nuanced, directly comparable with other exhibits, standardized and internally consistent. It can evidence the impact of exhibits on visitors and provide guidance for future development. Further discussion explores the benefit of these methods for museum practitioners and within the wider literature.

With a concern for lack of resources in conducting evaluation (Kubarek & Trainer, 2015), the MEUX evaluation methods produced an extensive amount and range of data using a short and minimally intrusive survey method that takes visitors on average 5 min to complete. Similarly,

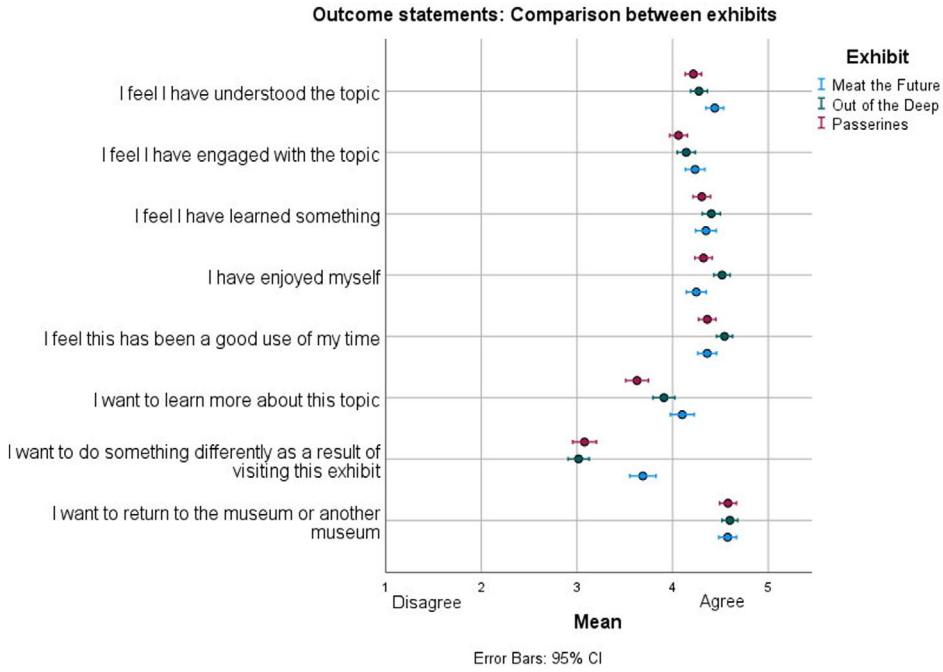


FIGURE 11 Comparison of outcome statements for three evaluated exhibits. [Colour figure can be viewed at wileyonlinelibrary.com]

the combination of qualitative and quantitative data approaches produces a rich and detailed understanding of a display or exhibition and the analysis uses evaluation results as evidence to support understanding, rather than presenting results in their own right. Instead of simply finding out that visitors felt they learned something from visiting the exhibit, we can also learn which elements of the exhibit influenced this directly and why visitors felt they were effective.

The MEUX framework provides a means to construct the visitor experience of an exhibit, with the ability to link inputs of the museum with visitor outcomes. There is the ability to understand what resources and efforts from the museum's part are having a definite effect on visitors. When considering this approach in relation to other evaluation frameworks within the museum sector, the benefits are threefold.

Reduction of positive bias

Current evaluation practices often suffer from a positive bias of results, where visitors report what the museum wishes to hear (Johanson & Glow, 2015). However, the use of the Semantic Differential scales, which are anchored at two points, provide parameters for visitors to evaluate an exhibit in that do not have an obviously "better" option. Therefore, visitors make more honest and objective judgments. For example, one of the lowest scoring scales for *Meat the Future* was "flat-dynamic." However, with COVID-19 restrictions in place, museum practitioners had made a conscious attempt to make the narrow exhibition gallery as open as possible, meaning much of the exhibit was flat against the walls, with less opportunities for touchables or interactives, which might have aided COVID transmission. Instead of being a negative, visitors had described what they saw, which had been a conscious decision by practitioners. This is therefore an improvement on current evaluation practice which has a tendency to provide overly positive results.

TABLE 10 Statistical results for Kruskal–Wallis one way analysis of variance of outcome statements between three different exhibits.

Outcome	Kruskal–Wallis test statistic	Difference
I feel I have understood the topic	0.000	Different
I feel I have engaged with the topic	0.004	Different
I feel I have learned something	0.103	Different
I have enjoyed myself	0.000	Different
I feel this has been a good use of my time	0.001	Different
I want to learn more about this topic	0.000	Different
I want to do something differently as a result of visiting the exhibit	0.000	Different
I want to return to the museum or another museum	0.958	Not different

Standardized results to support future development

Another limitation of existing evaluation practices is that results are contextualized only within the exhibit or program that they are evaluating and thus their generalisability or use is limited (Fu et al., 2016; Peterman et al., 2020; Teasdale, 2022). Where there is no one fixed approach to evaluation (Diamond et al., 2016) with evaluation methods evolving depending on the particular goals of an exhibit (Nelson & Cohn, 2015), the outcome is that the evaluation only ever provides specific data tied to an exhibit, rather than more general understandings of visitor experience that can be carried from exhibition to exhibition. However, it is this general understanding that allows for the implementation of recommendations and streamlining of practices, and so without them, it is evident that the long-term impact of evaluation is limited. In contrast, the use of the laddering interview technique elevates visitor opinions above the specific context of the exhibit and thus enables a deeper understanding of why a visitor finds an exhibit feature appealing that can be used to shape future practice.

For *Meat the Future*, visitors highlighted the scientific evidence being presented as an appealing feature. With the laddering technique, it was clear that the reason this was appealing was because it allowed visitors to feel empowered with information about an issue, which was important because it meant they could make their own decisions about their diet: a feature of the value of autonomy. With this method, a link is created between the specific context of the exhibit, and what this appeals to in visitors' belief and value systems. This provides us with a more general understanding of the appeal of the exhibition, which can then be implemented more easily in a future exhibit. Practitioners may still want to support visitor autonomy (the same value) but through a different method or topic.

The MEUX Evaluation Method therefore improves upon issues identified in the literature around evaluation results not being acted upon (Davies & Heath, 2013). Diamond et al. (2016) noted that "there is a gap between research and practice that can prevent evaluation results being acted upon," but as demonstrated with the MEUX Evaluation methods, the data collected are more detailed and useful for supporting future development. Thus, evaluation practices are more systematically integrated into development processes. Calls for more standardized approaches have been made in the recent literature (Ong & Ladenhead, 2015; Peterman et al., 2020; Voiklis et al., 2023) and it is clear that the MEUX Evaluation Method contributes to such efforts. This can be both in the temporary programs of one museum, but also apply across different museums and different exhibitions.

The study presented above was conducted at three different exhibits within one museum, OUMNH. These ranged from a traditional glass-case display to a larger, more modern exhibit with specimens, interpretation, and interactive elements, to a full temporary exhibition with multiple specimens, a longitudinal narrative, and various audio-visual and interactive elements. These exhibits were chosen to represent a range of exhibit styles and methodologies, and this highlights the benefit of the MEUX Evaluation Method in being able to draw direct comparisons across a variety of exhibit types. This extends further to being able to apply the method to different museums and, again, to draw direct comparisons. For example, research currently underway uses MEUX to objectively compare displays and exhibitions to three different UK museums. Using the standardized evaluation procedure, direct comparisons could be made between the science-based exhibits discussed in this study and the immersive history-based exhibits at the other museums, thus demonstrating how the MEUX Model and Evaluation Method is applicable to different exhibit and museum types across the UK sector.

Evaluation underpinned by conceptual framework

While there is extensive literature on exploring visitor behavior within exhibits and how visitors learn, engage and experience museum exhibits, these studies are often constrained by not being underpinned with a conceptual framework of visitor engagement. While constructivist theories of museum learning exist, they do not support practical research or understandings of such learning, and vice versa. For example, Falk and Dierking's Museum Experience framework (1992, 2016) sets out the Personal, Social, and Physical Contexts of museum learning, but there is no practical tool to measure these elements within museum settings. In reverse, extensive visitor research into different areas of the museum visit is not supported by a conceptual framework that allows individual results to contribute to wider, more general understandings of the visitor experience.

In more recent years, there have been some exceptions to this limitation, with increasing literature exploring the specific interactions between museums and their visitors and, on some occasions, how this is conceptualized theoretically (e.g., Baldioli et al., 2022; Ji et al., 2023; Packer et al., 2022; Krause & Davidson, 2021). One notable example of this is the work by Packer, Ballantyne and co-authors, which both conceptualizes the visitor experience theoretically (Packer et al., 2016), and provides an evaluation tool to measure such experience (Packer et al., 2018).

The DoVE Adjective Checklist (Packer et al., 2018) is based on the 10 facets of visitor experience (Packer et al., 2016). A total of 97 words or phrases were identified to represent these 10 facets, plus 19 negative phrases, for a total of 116. These were randomly listed and ordered on a single page with four columns, and visitors were asked to select all that they experienced more than they would in everyday life. There was also a reverse checklist. The problem with this method, however, is that it relies heavily on the order in which the terms are arranged, with a bias towards earlier listed items by up to 9% (Packer and Ballantyne, 2019), meaning that results rely in part on where a particular term is placed in the list. Furthermore, a tick-box selection of items limits the methods of statistical analysis, in that only frequencies of selections can be analyzed. While this does allow for highlighting significant differences in selections across sites, it is nevertheless limited in being unable to identify which elements of the exhibits or museum in general had an impact on visitors' selections. Therefore, while it provides a means to capture and evaluate visitor experience, Packer and Ballantyne's model (2016, 2019) is both subject to the ordering of the words for

evaluation (and thus less reliable) and lacking in understanding the relationship between the museum exhibit and the visitor outcome.

In general, the MEUX Model and Evaluation Method takes this recent trend of exploring interactions between museums and their visitors "at the exhibit face" in a more extensive and comprehensive way. The Model for MEUX (King et al., 2023) provides a nuanced and detailed conceptualisation of visitor experience. Using the evaluation methods described here, it has the power to capture the relationship between exhibits and visitors, understand how museum inputs can impact visitors, and provide more sophisticated statistical understanding of an exhibit. It is also possible to build a picture of a specific exhibit within the framework of the MEUX Model, and thus make comparisons between exhibits.

These standardized and generalized data means it is possible to utilize existing data in future exhibit development. It is possible to understand the deeper meanings for visitor appeal and lifted out of the context of a particular exhibit, we can implement these principles in other exhibits in the future. It does not necessarily place value judgments on different exhibits, with particular qualities being better than others, but rather allows us to select and make more conscious decisions about what qualities we would like an exhibit to have. If we would like an exhibit to be experientially led and inclusive for visitors, we can look to see how *Out of the Deep* has produced this visitor experience, but if in turn we would like visitors to be educated about an important issue, we can look to *Meat the Future* for guidance.

Limitations

Inevitably, there are some shortcomings to the MEUX method explored in this paper. The largest is perhaps the higher level language proficiency needed to take part in the survey. On some occasions, visitors expressed their confusion of some of the terms used in the survey. This is likely to cause difficult for those without English as a first language, or for children. Only visitors over the age of 18 took part in the studies, due to the nature of ethics approval, and so the survey has not been tested on children to date. This is an area for development for the research team to produce an equivalent survey which uses simpler terms or a method which is not in written form. Secondly, the data and techniques used in the MEUX method are relatively sophisticated, and perhaps difficult to implement if limited resources are available for evaluation, especially in smaller museums. To adopt this method in the sector more widely, further training and guidance is likely to be needed to ensure practitioners are empowered to use the MEUX tools successfully.

One other limitation may be the lack of specific questions that evaluate particular parts of an exhibition. For example, for *Meat the Future* this evaluation study did not specifically collect data on visitors' diets, despite this being a key factor in the engagement and learning potential for a visitor to the exhibition. There were several anecdotal pieces of evidence which suggested that audiences had a range of diets, including meat-eating and vegetarian or vegan, but this was not quantified and was not used as a factor in statistical analysis. We could not, for example, test whether a visitor's diet had an impact on their ratings of the exhibition, like we did for their gender or age group. However, the key benefit of the MEUX methods of evaluation is that they are standardized and allow for direct comparison between exhibitions. A question about diet would have therefore been inappropriate for visitors to *Out of the Deep* or *Passerines*. Furthermore, the researcher found that many visitors offered their diet choices freely when making other comments, as they themselves recognized its importance, and so some data was captured by these means. It may be the case that practitioners in future add specific questions for the exhibit they are evaluating, whilst maintaining the standardized approach for the rest of the questions to allow for comparability.

There are also a number of other elements of the visitor experience that are not directly captured by the MEUX Evaluation Method. These include the role of different visitors within a group, the amount of time spent at the exhibit, and pre-exhibit levels of knowledge and interest, which will all impact the visitor experience. However, it is not the case that the MEUX Evaluation Method can, or seeks to, capture every single detail of the visitor experience at an exhibit. It is comprehensive, but there is an inevitable trade off between the amount of detail captured, and the amount of resource used for such data capture, and thus it is argued that the MEUX Evaluation approach has a high yield of data for the resource required, even though this is not completely comprehensive.

In addition, while some elements of the visitor experience are not captured directly, there are a number of proxy indicators of these. An indication of a visitor's prior interest and knowledge is captured through their Motivation for Visiting, with educational motivation such a "I am interested in the topic" implying some level of prior knowledge and interest. Similarly, these educational motivations can also indicate how long a visitor spent in the exhibit, with those showcasing a stronger interest likely to also spend longer in the exhibit. Furthermore, cross-tabulation analysis allows us to identify whether this prior interest or exhibit behavior had a significant effect on the visitor experience and outcomes reported.

The role of different visitors within a group is also indicated by the question of Visit Group, where participants can report what type of group they were in. An important option on this which showcases visitor profiles is the option of "family with children," and it is therefore indicated that the adult who takes part in the survey on behalf of their family is acting in a group leader or facilitator type of role. Again, it is possible to test whether this role had an impact on the visitor experience through cross-tabulation analysis, and thus we can see how different group types behave differently in the exhibit.

Returning to the notion of a trade-off between detail and resource, the researchers made a conscious effort to produce an evaluation instrument that was detailed but not resource-heavy for museums to utilize in their daily work. While it is the case that more direct information about the visitor experience could be gained through visitor observations or pre-exhibit surveys, these would create additional resource for museums unnecessarily. Furthermore, there are wider questions on the validity of how observations or pre-exhibit surveys impact the visitor experience. The MEUX Evaluation Method is a post-exhibit survey and enables any visitor, regardless of the level of engagement they have shown, to take part if they wish. If participants were recruited in a pre-survey, this may impact their natural engagement with the exhibit, and thus may invalidate results. However, as discussed, the use of proxy indicators within the survey allows for many of these issues to be overcome.

Overall, the results presented demonstrate the benefits of using a User Experience approach and the MEUX Model in exhibit development and evaluation. The evaluation results highlight the varied nature of museum exhibits but also provide a means to standardize and directly compare results, meaning that knowledge of museum exhibit visitor experience can be built up over time. By grounding individual results within a consistent, repeatable framework, museum practices can become more efficient and build upon knowledge of how exhibits work that can provide benchmarking for future development.

CONCLUSION

Following on from research that conceptualized visitor experience in a Model for Museum Exhibit User Experience [King et al., 2023], this research has developed a number of evaluation methods that can be used to capture and assess visitor experience with different types of museum exhibits and exhibitions. Evaluation was conducted on three exhibits at [institution]

and results were explored in-depth with the *Meat the Future* exhibition. Qualitative and quantitative data together provided a picture of the exhibition utilizing the MEUX Model and detailed the relationship between museum inputs and how these affected and impacted visitor outcomes. These data were then used to provide direct comparisons between exhibits in a standardized and generalisable manner but did not necessarily provide a value judgment on which exhibit was 'better'.

Several benefits of the MEUX Model and evaluation methods are identified. Rich and detailed results were generated from a short and minimally intrusive visitor survey and interview of 5-min duration. Nevertheless, the analysis built a nuanced picture of the exhibits, with an understanding of how inputs from the museum as *Exhibit Features* related to how visitors perceived the *Exhibit Qualities* and how these related to *Visitor Gains*. Correlations provided further understanding of the mechanisms of visitor experience and which particular elements of an exhibit had the most impact on visitors. Using these data, it is possible to benchmark understanding of different visitor experiences and, in future exhibition development, to make conscious decisions about what an exhibit should be and how it should operate. This builds efficiency into exhibit practices by improving knowledge of what works, what does not work, and how museum resources are best used to make the most impact. In combining a theoretical framework and a practical means to assess this within and between institutions, the MEUX Model provides a vehicle to make evaluation practices within the museum sector more sophisticated, useful, and efficient, leading to more meaningful and impactful visitor experiences of exhibits and exhibitions.

ACKNOWLEDGMENTS

This work was supported by the Arts and Humanities Research Council under Grant AH/T002204/1.

CONFLICT OF INTEREST STATEMENT

There are no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Ellie King https://orcid.org/0000-0001-9573-1874
M. Paul Smith https://orcid.org/0000-0002-5141-1577
Paul F. Wilson https://orcid.org/0000-0002-7243-4372
Mark A. Williams https://orcid.org/0000-0003-3802-7914

REFERENCES

Alt, M B, Shaw, K M. 1984. Characteristics of Ideal Museum Exhibits. *British Journal of Psychology* 75. pp 25-36.

Angeli, Daniella D, Kelly, Ryan M, O'Neill, Eamonn. 2020. Beyond Happy-or-not: Using Emoji to capture visitors' emotional experience. *Curator: The Museum Journal* 63:2 pp 167–91.

Ardoin, Nicole M, Schuh, Janel S, Khalil, Kathayoon A. 2016. Environmental behaviour of visitors to a science museum. *Visitor Studies* 19:1 pp 77-95.

Asuero, A G, Sayago, A, González, A G. 2007. The Correlation Coefficient: An Overview. *Critical Reviews in Analytical Chemistry* 36:1. Pp 41-59.

Baker, Janice. 2015. Anarchical Artefacts: Museums as sites for radical otherness. In Witcomb, Andrea, Message, Kylie (eds.) *The International Handbooks of Museum Studies: Museum Theory*. John Wiley & Sons Ltd. Oxford. pp 63-77.

Baldioli, Shannon, Edson, Shauna, Grady, Ashley, Matos, Nefertiti. 2022. Touch Points: Co-designing tactile exhibition elements with user/experts. *Journal of Museum Education* 42:2 pp 166-78.

Bitgood, Stephen, Loomis, Ross J. 2012. Chan Screven's Contributions to Visitor Studies. *Curator: The Museum Journal* 55:2. pp 107-11.

Burmester, Michael, Mast, Marcus, Jäger, Kilian, Homans, Hendrik. 2010. Valence method for formative evaluation of user experience. *DIS '10: Proceedings of the 8th ACM Conference on Designing Interactive Systems*. Pp 364-7.

Coelho, Priscila Campos dos Santos, Rocha, Jessica Norberto, Massarani, Luisa. 2022. What do adolescents talk about when they visit an aquarium? A case study at the Marine Aquarium of Rio de Janeiro. *Visitor Studies* 25:1 pp 60-84.

Davies, M., Christian Heath 2013. Evaluating Evaluation: Increasing the Impact of summative evaluation in museums and galleries. visitors.org.uk.

Davis, Mark H, Highfill, Lauren E, Makecha, Radika, Baksic, William, Graves, Shane, Heaton, Emilia. 2023. Is there value in including information about animal cognition and emotion in zoo messaging? *Visitor Studies* 26:1 pp 42-58.

Del Chiappa, G., Andreu, L. and Gallarza, M.G. 2014. Emotions and visitor satisfaction at a museum. *International Journal of Culture, Tourism and Hospitality* 8:4 pp 420-31.

Diamond, Judy, Horn, Michael, Uttal, David H. 2016. *Practical Evaluation Guide: Tools for Museums and Other informal educational settings*. Rowman and Littlefield, London.

Falk, J. H., Lynn D. Dierking 1992. *The Museum Experience*, Washington, Whalesback Books.

Falk, JH, LD Dierking. 2016. *The Museum Experience Revisited*. Routledge, Oxon.

Falk, John H, Dierking, Lynn D. 2000. *Learning from Museums: Visitor experience and the making of meaning*. AltraMira. Plymouth.

Falk, John H, Dierking, Lynn D. 2018. *Learning from Museums*. Rowman and Littlefield. London.

Falk, John H. 2009. *Identity and the Museum Visitor Experience*. Routledge, New York.

Fu, Alice C, Kannan Archana, Shavelson, Richard J, Peterson, Lisa, Kurpis, Amy. 2016. Room for Rigor: Designs and methods in informal science education evaluation. *Visitor Studies* 19:1 pp 12-38.

Gardner, James B. 2015. Preserving/Shaping/Creating: Museums and public memory in a time of loss. In Witcomb, Andrea, Message, Kylie (eds.) *The International Handbooks of Museum Studies: Museum Theory*. John Wiley & Sons Ltd. Oxford. pp 511-30.

Griggs, Steven A. 1990. Perceptions of Traditional and New Style Exhibitions at the Natural History Museum, London, *ILVS Reviews* 1(2). pp 78-90.

Grunert, K. G., & Bech-Larsen, T. 2005. Explaining choice option attractiveness by beliefs elicited by the laddering method. *Journal of Economic Psychology*, 26:2, 223–41.

Harrasser, Karin. 2015. (Dis)playing the Museum: Artifacts, visitors, embodiment, and mediality. In Henning, Michelle (ed) *The International Handbooks of Museum Studies: Museum Media*. John Wiley & Sons Ltd. Oxford. pp 371-88.

Hassenzahl, M. 2003. The Thing and I: Understanding the Relationship Between User and Product. In: MA Blythe, A. M., K Overbeel, PC Wright (ed.) *Funology: From Usability to Enjoyment*. Netherlands: Kluwer Academic Publishers.

Hassenzahl, M., Effie Lou-Chong Law, Ebba Thera Huannberg 2008. User Experience (US): Towards an Experiential Perspective on Product Quality. In: E. Brangier, G. M., JMC

Hein, George. 1998. *Learning in the Museum*. Routledge, London.

Heuken, Nicole, Schuder, Anna-Lena, Christian, Andreas. 2021. Differences between first time and repeat visitors in a special exhibition at a Natural History Museum. *Visitor Studies* 24:1 pp 166-83.

Hoare, Jessica. 2020. The practice and potential of heritage emotion research: an experimental mixed-method approach to investigating affect and emotion in a historic house. *International Journal of Heritage Studies* 26:10 pp 955-74.

Hood, Marilyn. 1992. Significant Issues in Museum Audience Research. IVLS Review: A Journal of Visitor *Behaviour* pp 18-23.

Hooper-Greenhill, Eileen. 1992. *Museums and the Shaping of Knowledge*. Routledge. London.

Hooper-Greenhill, Eileen. 2000. *Museums and the Interpretation of Visual Culture*. Routledge. London.

Hooper-Greenhill, Eileen. 2007. *Museums and Education: Purpose, Pedagogy, Performance*, Routledge. Oxon.

Izard, C. 1979. The Differential Emotions Scale. Unpublished manuscript, University of Delaware.

Ji, Jiao, Anderson, David, Wu, Xinchun. 2023. Chinese science museum educators' perspectives on the contradictions between their practice and visitors' expectations. *Visitor Studies* 26:2 pp 1-21.

Johanson, Katya, and Glow, Hilary. 2015. A Virtuous Circle: The positive evaluation phenomenon in arts audience research. *Participations: Journal of Audience and Reception Studies* 12. pp 254-70.

Johnston, David. 1999. *Assessing the Visiting Public's Perceptions of the Outcomes of Their Visit to Interactive Science and Technology Centres*. Curtin University. http://hdl.handle.net/20.500.11937/61

Karlsson, M.A., Wikstrom, L., 1999. *Beyond aesthetics! Competitor advantage by an holistic approach to product design, sixth International Product Development Management Conference*, Cambridge, UK, pp. 629–38.

King, E., Smith, M. P., Wilson, P. F., Stott, J., & Williams, M. A. (2023). Creating Meaningful Museums: A Model for Museum Exhibition User Experience. *Visitor Studies*, 26(1), 59–81.

Krause, Amanda E, Davidson, Jane W. 2021. Investigating the development and reception of an art exhibition on the theme of Early Modern representations of love. *Curator: The Museum Journal* 64:4 pp 693-714.

Kubarek, Joy. Trainer, Laureen. 2015. Empowering Museum Educators to Evaluate. *Journal of Museum Education* 40:1. pp 3-7.

Laviea, Talia, Tractinsky, Noam. 2004. Assessing dimensions of perceived visual aesthetics of websites. *International Journal of Human-Computer Studies* 60. pp 269-98.

Liikkanen, Lassi A, Reavey, Heather. 2015. Resonance testing: an industry approach for experiential concept evaluation. *International Journal of Product Development*, Inderscience Enterprises Ltd. 20:4, pp 265-85.

Lindauer, Margaret. 2005. What to ask and how to answer: A comparative analysis of methodologies and philosophies of summative exhibit evaluation. *Museum and Society* 3. pp 137-52.

Lucija, Andre, Curksen, Tracy, Volman, Monique L. 2017. Museums as avenues of learning for children: a decade of research. *Learning Environments Research* 20 pp 47-76.

Luebke, Jerry F. 2018. Zoo exhibit experiences and visitors' affective reactions: a preliminary study. *Curator: The Museum Journal* 61:2 pp 345-52.

Macdonald, Sharon. 2006. Expanding Museum Studies: An Introduction. In Macdonald, Sharon (ed.) *A companion to museum studies*. Blackwell. Oxford. pp 1-12.

MacPherson, Anna, Hammerness, Karen, Gupta, Preeti. 2019. Developing a set of guidelines for rigorous evaluations at a Natural History Museum. *Journal of Museum Education* 44:3 pp 277-85.

May, Sarah, Todd, Katie, Daley, Samantha G, Rappolt-Schilchtmann, Gabrielle. 2022. Measurement of science museum visitors' emotional experiences at exhibits designed to encourage productive struggle. *Curator: The Museum Journal* 65:1 pp 161-85.

Miles, R., Alan Tout 1994. Impact of Research on the Approach to the Visiting Public at the Natural History Museum, London. In: Hooper-Greenhill, E. (ed.) *The Educational Role of the Museum*. London: Routledge.

Munro, Ealasaid. 2014. Doing emotion work in museums: reconceptualising the role of community engagement practitioners. *Museum and Society* 12:1 pp 44-60.

Nelson, Amy, and Cohn, Sarah. 2015. Data Collection Methods for Evaluating Museum Programs and Exhibitions. *Journal of Museum Education* 40:1. pp. 27-36.

Ntamkarelou, Lydia, Bantimarovdis, Philemon, Economou, Maria. 2017. Testing the uses and gratifications approach to museum visiting: adopting a mediated perspective in the cultural domain. *Visitor Studies* 20:1 pp 56-71.

O'Connor, Molly C, Nelson, Kristen C, Pradhananga, Anuit, Earnest, Megan E. 2020. Exploring how awareness-making elicits meaning-making in museum visitors: A mixed-method's study. *Journal of Museum Education* 45:2 pp 187-99.

Ong, Angelina, Ladenhead, Christina. 2015. Using one evaluation to affect many. *Journal of Museum Education* 40:1 pp 55-61.

Osgood, Charles E. 1964. Semantic Differential Technique in the Comparative Study of Cultures. *American Anthropologist* 66:3:2. pp 171-200.

Packer, Jan, Ballantyne, Roy, Bond, Nigel. 2018. Developing an Instrument to capture multifaceted visitor experiences: The DoVE Adjective checklist. *Visitor Studies* 21:12. pp 211-31.

Packer, Jan, Ballantyne, Roy, Hughes, Karen, Sneddon, Joanna, Lee, Julie. 2022. Differences between zoo/aquarium staff and visitors' values, beliefs and pro-environmental behaviours: consequences for environmental communication. *Visitor Studies* 25:1 pp 85-103.

Packer, Jan, Ballantyne, Roy, Luebke, Jerry F. 2016. Conceptualising the Visitor Experience: A review of literature and development of a multifaceted model. *Visitor Studies* 19:2. pp 128-43.

Pecore, John L, Kirchgesseur, Mandy L, Demetrikopoulos, Melissa K, Carruth, Laura K, Frantz, Kyle J. 2017. Formal lessons improve informal educational experiences: the influence of prior knowledge on student engagement. *Visitor Studies* 20:1 pp 89-104.

Pekarik, Andrew J, Doering, Zahava D, Karns, David A. 1999. Exploring Satisfying Experiences in Museums. Curator: The *Museums Journal* 42:2. Pp 152-73.

Peterman, Karen, Verbeke, Monae, Nielsen, Katherine. 2020. Looking back to think ahead: reflections on science festival evaluation and research. *Visitor Studies* 23:2 pp 205-17.

Phelan, Sielle, Specht, Inga, Lewalter, Doris. 2020. Visitor motivation as part of visitors' personal context in a science museum. *Visitor Studies* 23:2 pp 141-61.

Pietila, Misa. 2017. Do visitor experiences differ across recreation settings? Using geographical information systems to study the setting-experience relationship. *Visitor Studies* 20:2 pp 187-201.

Price, Aaron C, Applebaum, Lauren. 2022. Measuring a sense of belonging at museums and cultural centres. *Curator: The Museum Journal* 65:1 pp 135-60.

Price, Aaron C, Greenslit, Jana Nicole, Applebaum, Lauren, Harris, Natalie, Segovia, Gloria, Quinn, Kimberly A, Krogh-Jespersen, Sheila. 2021. Awe and memories of learning in science and art museums. *Visitor Studies* 24:2 pp 137-65.

Radywyl, Natalia, Barikin, Amelia, Papastergiadis, Nikos, McQuires, Scott. 2015. Ambient Aesthetics: altered subjectivities in the new museum. In Witcomb, Andrea, Message, Kylie (eds.) *The International Handbooks of Museum Studies: Museum Theory*. John Wiley & Sons Ltd. Oxford. pp 417-36.

Rennie, Leonie J, Johnstone, David J. 2004. The nature of learning and its implications for research on learning from museums. *Science Education* 88 pp S4-S16.

Rennie, Leonie J, Williams, Gina F. 2002. Science centers and scientific literacy: Promoting a relationship with science. *Science Education* 86:5. pp 706-26.

Reynolds, Thomas J, Olson, Jerry C. (eds.) 2001. *Understanding Consumer Decision Making: The means-end approach to marketing and advertising strategy*. Psychology Press. New York.

Roberts, Lisa C. 1997. From Knowledge to Narrative: Educators and the Changing Museum. *Visitor Behaviour* 7. Pp 7-8.

Robertson, Jane Evia, Peterman, Karen. 2020. Understanding engagement with science festivals: who are the engaged? *Visitor Studies* 23:1 pp 66-81.

Robinson, Cyntha. 2021. Museums and Emotions. *Journal of Museum Education* 46:2 pp 147-9.

Rodehn, Cecilia. 2020. Emotions in the museum of medicine. An investigation of how museum educators employ emotions and what these emotions. *International Journal of Heritage Studies* 26:2 pp 201-13.

Roschelle, Jeremy. 1995. Learning in interactive environments: prior knowledge and new experience. In J. H. Falk & L. D. Dierking (eds.) *Public Institutions for Personal Learning: Establishing a Research Agenda*. American Associations of Museums. Washington. Pp 37-51.

Roseman, Ira J. 1996. Appraisal Determinants of Emotions: Constructing a More Accurate and Comprehensive Theory. *Cognition and Emotion* 10:3 pp 241-78.

Scherer, K. R. 2005. What are emotions? And how can they be measured? *Social Science Information*, 44:4. pp 693-727.

Schorch, Philipp. 2015. Museum Encounters and narrative engagements. In Witcomb, Andrea, Message, Kylie (eds.) *The International Handbooks of Museum Studies: Museum Theory*. John Wiley & Sons Ltd. Oxford. pp 437-57.

Shaby, Neta, Ben-Zui, Orit, Tal, Assarat Tali. 2019. Engagement in a science museums – the role of social interactions. *Visitor Studies* 22:2 pp 1-20.

Sneddon, J.N, Hughes, K, Ballantyne, R. 2021. Using values-based focus groups to evaluate visitors' perceptions and interpretation of a wildlife exhibit. *Visitor Studies* 24:1 pp 100-120.

Teasdale, Rebecca M. 2022. How do you define success? Evaluative criteria for informal STEM education. *Visitor Studies* 25:2 pp 163-84.

Trucks, Jessica, Schmoll, Shannon, Hinko, Kathleen, Lopez, Gloria. 2022. How interested are planetarium visitors in astronomy? Comparing online and in-person audiences. *Journal of Museum Education* 47:4 pp 510-520.

Van Kleef, E, van Trijp, H, Luning, P. 2005. Consumer research in the early stages of new product development: A critical review of methods and techniques. *Food Quality and Preference* 16. pp. 181-201

Voiklis, John, Flinner, Kate, Field, Shaun, Gupta, Rupanwita, Fraser, John, Dwyer, Joseph de la Torre, Rank, Shelley, Nock Katherine. 2023. Seeing the forest not the trees: crowdsourced data collection methods for sector-wide research. *Visitor Studies* 26:1 pp 24-41.

Voss, K. E., Spangenberg, E. R., & Grohmann, B. (2003). Measuring the Hedonic and Utilitarian Dimensions of Consumer Attitude. *Journal of Marketing Research*, 40(3), 310–320.

Watson, David, Clark, Lee Anna. 1994. *The PANAS-X: Manual for the Positive and Negative Affect Schedule – Expanded Form*. University of Iowa.

Watson, Sheila. 2015. Emotions in the History Museum. In Witcome, Andrea, Message, Kylie (eds.) *The International Handbooks of Museum Studies: Museum Theory*. John Wiley & Sons Ltd. Oxford. pp 283-301.

Wellings, Tom, Williams, Mark A., Pitts, Matthew. 2008. Customer perception of switch-feel in luxury sports utility vehicles. *Food Quality and Preference* 19:8 pp 737-46.

Zaman, Bieke. 2008. Introducing contextual laddering to evaluate the likeability of games with children. *Cognition, Technology & Work* 10:2. pp. 107-17.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

AUTHOR BIOGRAPHY

Ellie King is an AHRC funded Collaborative Doctoral Programme PhD student researching exhibition development and user experience at Warwick Manufacturing Group, University of Warwick and Oxford University Museum of Natural History. Ellie's PhD is supervised by the co-authors of this paper.

Professor **M. Paul Smith** is the Director of Oxford University Museum of Natural History. Paul has interests in the application of digital technologies to science museums, particularly in the areas of 3D visualisation, virtual reality and the evaluation of user experience.

Dr Paul F. Wilson is a research fellow at Warwick Manufacturing Group, University of Warwick. His research focuses on 3D visualisation within museology and the cultural heritage sectors.

Janet Stott is the Deputy Director and Head of Public Engagement at Oxford University Museum of Natural History.

Professor **Mark A. Williams** is Leader of the Centre for imaging, Metrology, and Additive Technology (CiMAT) at Warwick Manufacturing Group, University of Warwick. Within this diverse role, Mark has an interest in digital conservation in the heritage and archaeology sectors.

How to cite this article: King, Ellie, M. Paul Smith, Paul F. Wilson, Janet F. Stott and Mark A. Williams 2025. "Evaluating Museum Exhibits: Quantifying Visitor Experience and Museum Impact With User Experience Methodologies." *Curator: The Museum Journal* 68 (1): 163–200. https://doi.org/10.1111/cura.12637.