

Research Article

Volker Eisenlauer*, Heinrich Söbke, Bernhard Ertl, Maximilian Fink and Florian Wehking

Educational virtual field trips 2035

The evolution and prospects of immersive learning through virtual explorations

<https://doi.org/10.1515/icom-2026-0010>

Received February 25, 2026; accepted February 25, 2026;

published online April 7, 2026

Abstract: Virtual field trips (VFTs) are learning activities offering unique learning affordances, such as fostering object-based learning through interaction with digital artifacts, as well as supporting experiential and situated learning by placing students in contexts that simulate real-world experiences. 2D virtual field trips have been around for several decades, with their formats continuously evolving in response to new technological advancements. Initially dominated by images, graphics, and videos, VFTs evolved significantly with the introduction of 360° media, game engines and the metaverse, which enabled more immersive, interactive, and collaborative experiences. In this article, we outline evidence-based development paths and potential visions for the future of VFTs over the next 10 years. A retrospective of VFTs reveals their evolution from simple image and video-based experiences to increasingly immersive and interactive formats. The specific features of state-of-the-art VFTs can be identified using the concept of affordances which refer to the action possibilities provided by underlying

technologies, such as 360° media, Computer Generated Imagery (CGI)-based VR worlds, and photogrammetry. The specification of action provided by current immersive technologies serves as a starting point into understanding future developments of VFTs. As part of this process, 14 hypotheses on the future development of VFTs were formulated, subsequently evaluated by experts in a Delphi study and then edited based on this expert feedback, providing empirical insights into expected trends and future directions for VFTs. The identified capabilities and educational potential will form the basis for expert discussions on upcoming trends in VFTs.

Keywords: virtual field trips; 360° media; educational technologies; affordances; immersive learning; Delphi study

1 Introduction

VFTs have their roots in traditional field trips (FTs), which have long been valued for their ability to provide experiential learning opportunities. Knutson¹ defines FTs as “shared social experience that provides the opportunity for students to encounter and explore novel things in an authentic setting”. FTs enable experience-based learning and have been shown to be significantly more conducive to learning than a comparable classroom lesson,² for example by increasing motivation and interest in the location visited.³ The advantages of FTs are particularly evident when they are not simply a substitute for a lecture, but are designed according to their strengths and integrated into a didactic setting as opportunities for exploration, discovery and learning experiences.⁴ Despite their high didactic merit, FTs also have some disadvantages that may well prevent their use. These include the time and effort required to conduct a FT. Inaccessibility of the locations in question may also prevent FTs. As an alternative to FTs, virtual field trips (VFTs) (or virtual tours) have been discussed for several decades. A common definition of a VFT is “a journey taken without actually making a trip to the site”.⁵ Woerner

*Corresponding author: Volker Eisenlauer, Institut für Bildungswissenschaft, Universität der Bundeswehr München, Neubiberg, Germany, E-mail: volker.eisenlauer@unibw.de.

<https://orcid.org/0000-0002-0710-430X>

Heinrich Söbke, Professor für Interaktive Digitale Systeme, Hochschule Weserbergland, Hamelin, Germany; and Bauhaus-Universität Weimar, Bauhaus-Institut für zukunftsweisende Infrastruktursysteme (b.is), Weimar, Germany, E-mail: soebke@hsw-hamelin.de.

<https://orcid.org/0000-0002-0105-3126>

Bernhard Ertl and Maximilian Fink, Professur für Lernen und Lehren mit Medien, Institut für Bildungswissenschaft, Universität der Bundeswehr München, Neubiberg, Germany,

E-mail: bernhard.ertl@unibw.de (B. Ertl),

maximilian.fink@yahoo.com (M. Fink)

Florian Wehking, Bauhaus-Universität Weimar, Bauhaus-Institut für zukunftsweisende Infrastruktursysteme (b.is), Weimar, Germany,

E-mail: florian.wehking@uni-weimar.de

further explains that such trips are made using alternative means, including media such as slides, videos or websites, panoramic photos, hypertext or even 3D models. In addition, we hold the exploratory nature of VFTs a defining feature. From our review of various VR learning environments, VFTs consistently prioritize self-directed exploration, as opposed to other environments that focus, for example, on the training of procedural knowledge and motor skills.

Overall, VFTs aim to convey the experience of being in a real place without physically being there. This can be achieved through various methods, including physical replicas of key features of the location or, more commonly, through photos and videos that depict places and objects in detail. As digital technology advances, physical items may be replaced by digital artifacts such as photographs,⁶ 360° (panoramic) photographs, hypertext^{7,8} or CGI models.⁹ Currently, 360° technology is one of the most utilized tools in various educational contexts, valued for its ease of access and flexible implementation.^{10–15} Among the disciplines in which VFTs are used are civil and environmental engineering,^{16,17} history,¹⁸ natural history,¹⁹ paleontology²⁰ and social work.²¹ Furthermore, 360° media continue to advance with new learning-enhancing features, such as multi-user functionality.^{22–24}

In line with this, our study seeks to address the following research questions: 1) How is the development of VFTs linked to technological advancements? 2) What affordances do VFTs provide from the perspectives of technology, experience, and learning? 3) How are VFTs expected to evolve in the future?

Conceptually, we use Gibson's²⁵ notion of *affordances* to link technological features of VFTs to the experiences they enable and the learning activities they support. Within the experiential dimension, we further specify spatial experience through Zhao & Klippel's²⁶ concepts of *place* and *scale*, which captures how learners perceive meaningful locations and perspective shifts in VFTs. Methodologically, the study combines an affordance-based theoretical lens with validation through a Delphi survey. The theoretical part defines key features and affordances of state-of-the-art VFTs considering their technological foundations, experiential qualities, and educational potential. From this analysis, 14 well-founded hypotheses were derived to anticipate how VFTs may evolve in terms of technology, learning experience, and organizational design. In the empirical part, these hypotheses were evaluated by 24 experts using a simplified Delphi method,²⁷ yielding empirically grounded insights into emerging trends and likely directions of VFTs over the next 10 years.

2 State of the art – VFT technologies, experiences, and learning

The concept of affordances, originally introduced by Gibson,²⁵ offers a useful way to describe how technological innovations translate into particular forms of experience and, potentially, learning. Affordances refer to the “action possibilities” that an environment provides to an individual. Such possibilities depend not only on the properties of the environment but also on the user's abilities and prior knowledge as well as on situational conditions. In the context of virtual field trips, this highlights that design and media decisions shape what learners can perceive and do: for instance, whether a VFT relies on 360° video, photogrammetry, or CGI influences interactivity (i.e., what actions are possible), and the kinds of exploration that are feasible. Building on this understanding, the following analysis uses an affordance-based lens to develop propositions about future developments in VFTs.

- Technological affordances describe system-level action possibilities enabled by production, post-production, and reception technologies – including the interaction techniques and devices through which users can act in the virtual environment.
- Experiential affordances describe how these technological conditions shape user perception and spatial experience. Immersion is treated here as an experiential effect (i.e., the felt sense of “being there”) that can emerge from interaction and sensory presentation. To further specify spatial experience, we draw on Zhao & Klippel's²⁶ concepts of place and scale as two concrete facets of spatial experience and perspective-taking.
- Learning affordances describe learning-promoting activities and outcomes that can be supported when technological and experiential affordances are aligned with instructional aims.

The three perspectives are analytically distinct but connected. The same design feature can be described at different levels: for example, “hotspots” are a technological affordance (selection/navigation), they shape experiential affordances (agency and guided attention in a place), and they support learning affordances when embedded into tasks (e.g., inquiry prompts, comparisons). In other words, technological affordances enable forms of (inter-)action, experiential affordances describe how these interactions and presentations are experienced (including immersion,

place, and scale), and learning affordances describe how these conditions can be used pedagogically.

Technological Affordances. A VFT created using CGI allows for dynamic, fully interactive environments, where learners can explore and manipulate objects, while 360° media offer more limited interactivity, typically confined to navigating the depicted environment and clicking on hotspots that activate embedded images, text, videos, or audio. These production technologies differ not only in representational format (photo-based vs. model-based), but also in the degree to which learners can move beyond observation toward interaction: 360° media mainly support viewpoint changes and information access, whereas photogrammetry adds navigable 3D inspection of captured objects and spaces, and CGI enables rule-based interaction (e.g., manipulation, simulation, multi-user actions). Post-production software also plays a crucial role in shaping the affordances of VFTs, as it determines how effectively the content can be presented and brought to life (e.g., hotspots, branching, embedded media) and how interaction is orchestrated. Here, authoring tools typically provide predefined interaction (e.g., linking scenes, hotspot logic), whereas game engines allow more complex interaction mechanics (e.g., physics, conditional events, adaptive

feedback) at the cost of higher development expertise. Likewise, reception technology constrains what kinds of actions are possible: a VFT can be viewed on a computer screen, through a head-mounted display (HMD) or in a VR cave – each of which affords different interaction techniques and degrees of embodied control. Drawing on research on multimedia/VR and interactivity,^{28,29} VFT interactivity can be differentiated into structural (navigation, hotspots, branching – common in 360° VFTs), cognitive (interaction features that scaffold sensemaking, e.g., cues/questions and comparisons), and productive (creating/modifying artifacts or manipulating objects – more typical in CGI). A further technological dimension is the means of interaction (e.g., mouse/keyboard/touch, controllers, hand tracking/gloves, full-body/camera-based interaction), which shapes what kinds of structural, cognitive, and productive interactivity can be implemented and at what level of embodied control. Table 1 summarizes the key affordances across production, post-production, reception and means of interaction technologies, highlighting how they shape the forms and (inter) action possibilities in VFTs.

Experiential Affordances. VFTs provide immersive experiences, enabling learners to explore and interact with environments either captured from the physical world or

Table 1: Key affordances of production and interaction technologies in virtual field trips.

System level	Type	Affordances
Production	360° media	Captures real-world environments; viewpoint navigation; hotspot-based access to embedded media; limited object manipulation
	Photogrammetry	Creates navigable 3D replicas; spatial inspection of detailed replicas; limited-to-moderate object interaction depending on implementation
	CGI	Fully interactive environments; rule-based object manipulation; simulation; multi-user interaction (depending on system design)
Post-Production	Authoring tools	Template-based interaction logic (e.g., navigation, hotspots, branching, embedded media) with standardized workflows
	Game engines	Highly dynamic interaction mechanics (e.g., physics, conditional events, adaptive logic); supports complex simulations and multi-user functions
Reception	WebVR	Screen-based viewing; efficient navigation and information access; interaction typically via standard desktop/mobile inputs
	Head-mounted displays	Head-tracked viewing; embodied viewpoint control; supports more natural interaction techniques depending on hardware/software
	VR caves	Shared, co-located 3D viewing; tracked interaction in space; supports group exploration and coordinated interaction
Means of interaction	Mouse/keyboard/touch	Pointing & selection; menu-based navigation; hotspot activation; limited object manipulation
	Controllers (6 Degrees of freedom, 6DoF)	6DoF pointing and selection; grabbing/manipulation in 3D space; support for tool-like actions (e.g., rotate/scale)
	Hand tracking/gloves	Direct hand-based grabbing and gestures; more natural object manipulation; fine-grained interaction with virtual tools and interfaces
	Full-body	Whole-body movement and spatial interaction; embodied navigation; posture/locomotion-based interaction and co-located collaboration

fully created through digital technologies. These experiences vary in immersion, interactivity, and agency, shaping how learners perceive and navigate virtual environments. To specify experiential affordances more precisely, we use the concepts of place and scale²⁶ as organizing sub-dimensions of spatial experience in VFTs. Importantly, place and scale are not mutually exclusive; both can co-occur in the same VFT, with emphasis shifting depending on learning goals. “Place” is broadly defined as a location imbued with meaning, which immersive technologies allow learners to experience in meaningful ways. Emerging technologies like 360° media enable the exploration of various locations without physically visiting them. Moreover, virtual places can extend beyond reality, incorporating interactive elements like clickable objects, enhancing the educational experience by enabling learners to engage with the content in otherwise inaccessible, exploratory ways. According to Zhao & Klippel²⁶ “scale” refers to the perceived size of a place in relation to the human body, providing a sense of how large or small a place appears from the viewer’s perspective. Via and within immersive VFTs places can be easily presented at varying scales. The concepts of “place” and “scale” can be vividly illustrated in the VFT to the Paleontological Collection in Munich: A close-up view of the ancient elephant skeleton represents “place” as a meaningful location. Through 360° media, learners can examine the skeleton’s details up close, offering an experience that would not be possible in person, as the fossil is typically viewed from behind barriers (Image 1). Similarly, a bird’s-eye perspective, such as viewing the museum foyer and the ancient elephant skeleton from a vantage point near the pterosaur, illustrates the concept of “scale” (Image 2).

From this vantage point, the viewer experiences the vastness of the space, gaining a sense of how the skeleton fits within the larger environment. This highlights how VFTs allow for dynamic shifts in perspective, helping learners understand the relationship between their own scale and different objects of the surrounding space. Moreover, VFTs



Image 1: “Place” – ancient elephant skeleton in detail.



Image 2: “Scale” – foyer and skeleton from above.

may provide access to inaccessible locations, such as waterworks³⁰ or perspectives, e.g., by drone flights.³¹ In a recent study, students learning about wastewater treatment via a VFT improved their test scores by 20%.³² The authors highlight that exploring different stages of the treatment process in a virtual environment made complex concepts more accessible, while also eliminating the logistical challenges of organizing traditional FTs.

Learning Affordances. From a learning perspective, learning affordances emerge when technological affordances (i.e., the available inter-action possibilities) and experiential affordances (i.e., immersion, agency, and spatial experience) are aligned with instructional aims. While more interactive and immersive technologies may offer deeper engagement, they do not necessarily lead to better retention.³³ In fact, overly stimulating environments can result in focus impairment and cognitive overload, potentially hindering learning.³⁴ Therefore, when planning and designing VFTs, it is crucial to align learning objectives with the technological and experiential conditions a VFT affords, ensuring that both content creation and content reception are carefully considered. For teachers, this alignment also involves a pragmatic choice of production approach: 360°-based VFTs typically offer a low barrier to entry and can be authored without coding, whereas CGI-based VFTs enable richer interaction but usually require specialized expertise and higher development effort. From the literature, VFTs are attributed various learning-promoting affordances, including: 1) *Self-directed learning*. The option of learning independently of time and (almost) location enables learners to arrange their own learning process. Minor restrictions on location independence can result from the hardware used, for example through stationary HMDs. 2) *Active Learning*. VFTs can support active learning when learners actively investigate content through exploration, comparison, and hypothesis testing. This draws on cognitive interactivity and is enabled by structural interactivity (navigation/choices); productive actions such as

annotation or object manipulation can further deepen engagement when integrated into tasks. As a result, VFTs can enable exploration and experimentation that would not be possible in real field trips due to real-world constraints.

3) *Explorative Learning*. VFTs allow learners to explore environments from different perspectives and understand spatial relationships in a more meaningful way. This dynamic interaction helps them to bridge abstract concepts and real-world applications, enhancing engagement and retention.³¹

4) *Situated Learning*. VFTs allow for situated learning in alignment with situated experiential education,¹⁷ where learning occurs most effectively when learners can engage with content in a relevant context.³⁵

Taken together, this affordance-based perspective directly addresses the three research questions mentioned above concerning the development of VFTs in relation to technological progress, the learning experience and the general future of VFTs. Although social and intercultural, affective, and ethical issues also play a central role in VFTs, they are less well-suited to being framed as generalizable dimensions, as they are strongly shaped by the specific content and contextual embedding of a given VFT. It is therefore recommended that they be examined separately in future work.

3 Future VFTs—technological, didactic, and organizational prospects

The evolution of VFTs underscores the close interplay between technological innovation, the immersive experiences it enables, and their resulting educational value. As outlined in Chapter 2, affordances depend not only on system properties but also on users' prior knowledge and situational conditions. While these user- and context-specific factors are beyond the scope of this paper, we address the

contextual side by considering the organizational conditions that shape how VFTs are produced, distributed, and sustainably used in educational practice. As a foundation for our Delphi study, we therefore structured prospective developments along the dimensions of technology, experience, and education, complemented by an organizational perspective. While Chapter 2 distinguishes experiential and learning affordances analytically, we discuss them jointly under didactical prospects here, because didactic use links how VFTs are experienced (e.g., immersion, place/scale) to the learning activities they can support. Building on Chapter 2, we first identified core technological, experiential, and learning affordances of current VFTs (i.e., action possibilities enabled by production/post-production/reception technologies and their experiential and pedagogical implications). We then related these affordance patterns to current technological and institutional trajectories (e.g., AI-driven interaction, mixed reality hardware, digital twins, learning analytics, social VR, and platformization) and translated the resulting directions into concrete, testable statements about likely developments by 2035 (H1–H14). Figure 1 provides a schematic overview of this process, from affordance patterns to trajectories and hypotheses (H1–H14). Each hypothesis is assigned to the type of prospect it mainly concerns, while acknowledging overlap across prospective VFT affordances.

Technical Prospects (TP). Our anticipated technical prospects build directly upon the *technological affordances* previously discussed and incorporate emerging innovations that are already shaping the use of VFTs. Future VFTs will increasingly integrate diverse technologies, creating richer, more interactive, and adaptable experiences. A key development is the fusion of 360° imagery with CGI, enabling dynamic virtual landscapes that combine realism with simulation. Game engine-based VFTs using tools like Leap Motion already demonstrate this potential,³⁶ where the presentation of the tracked hands is achieved through

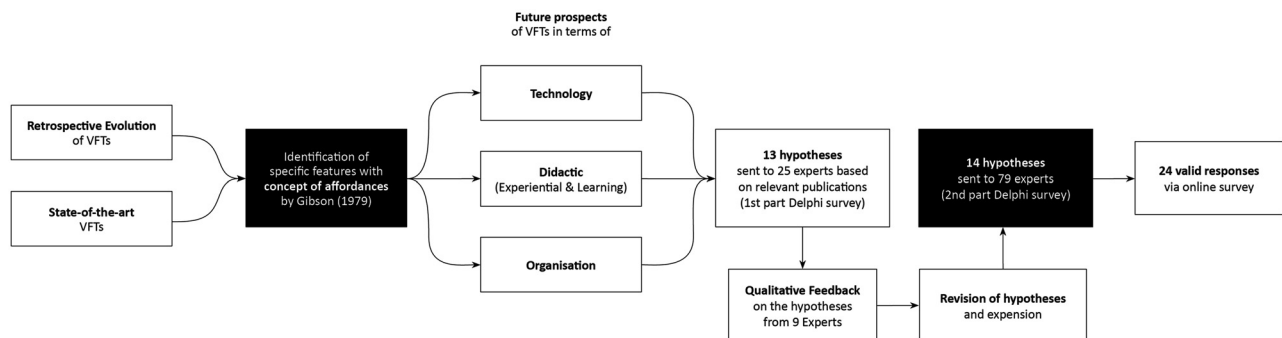


Figure 1: Hypothesis derivation process.

Unity-based CGI, merging precise motion tracking with realistic textures for an immersive experience [see H1 – *Design and Production*]. AI-powered avatars represent humans authentically with videos or 3D models and are powered by large language models and other machine learning algorithms.³⁷ They will facilitate the embedding of human-like interaction partners in VFTs and may be used to deliver personalized guidance and contextualized learning [H4 – *Interactive Avatars*]. Sensors in digital twins provide precise data-driven analysis, while 360° VFTs enable visual exploration of structures. Combining both links real-time physical data with immersive visual context, enabling virtual scenarios to be represented more realistically and responsively [H5 – *Physical Data*]. 360° cameras facilitate live-streamed VFTs, bridging virtual and physical field trips [H6 – *Immersive Live Streaming*]. Lastly, mobile XR headsets will likely replace traditional screens as the dominant access point for VFTs. These devices offer flexible, immersive engagement and are expected to become the norm by 2035 [H3 – *Reception*].

Didactical prospects (DP). The didactical potential of VFTs lies in their capacity to engage learners through authentic, interactive, and situated experiences. They foster self-directed, active, and explorative learning by enabling perspective shifts, object interaction, and access to otherwise inaccessible spaces. Looking ahead, future developments are expected to further enhance these qualities – particularly through adaptive, high-resolution visuals that respond to learners’ needs. In this context, the concepts of “place” and “scale”²⁶ will become increasingly relevant, as dynamic representations deepen spatial understanding and strengthen connections between abstract concepts and real-world phenomena [H10 – *Adaptive Visual Representations*]. Advances in learning analytics further support personalized learning experiences. By tracking user behaviour, systems can identify areas of difficulty and deliver adaptive content, helping to maintain engagement and improve retention [H9 – *Learning Analytics*]. For VFTs to have a lasting impact, however, their didactical integration must go beyond individual experiences. By 2035, social VR platforms will enhance multi-user interactions in shared environments³⁸ and thus enable collaborative learning [H12 – *Collaborative Learning*]. At the same time, inclusive design principles will be essential to ensure that VFTs accommodate the diverse physical, cognitive, and social needs of all learners – enabling truly accessible and equitable immersive education [H14 – *Accessibility*].

Organizational Prospects (OP). To ensure VFTs reach their full didactical impact, they must be embedded within

supportive organizational structures. User-friendly authoring tools will lower barriers to VFT creation. By 2035, teachers and students alike can craft custom virtual environments, fostering creativity and deeper engagement with the subject matter [H2 – *Content Preparation*]. The advent of AI-based development tools could further streamline this process, reducing the need for technical expertise and enabling broader participation in content creation [H8 – *Development Tools*]. In analogy to established platforms, such as MOOC platforms, centralized repositories will provide structured, searchable access to modular VFTs, which will facilitate curriculum integration and promote consistent use across institutions and disciplines [H7 – *Delivery*]. Intelligent recommender systems will suggest VFTs based on user interests, learning history, and curricular goals, thereby fostering autonomous exploration and sustained motivation [H11 – *Recommender-Based Experiences*]. For lasting impact, VFTs must be embedded in teacher education and professional development. Educators will need competencies not only in applying VFTs, but also in adapting them to diverse pedagogical contexts. By 2035, such training is expected to be a standard part of teacher qualification [see H13 – *Integration into Teacher Education*].

4 Delphi survey

The developed hypotheses were subjected to a two-round Delphi survey. Experts were selected based on relevant publications in the fields of VFTs, educational 360° media, and educational VR, and were contacted via email. In the first round, nine experts – Table A-1 shows their demography – provided qualitative feedback on the initial set of hypotheses, which had been developed during a workshop held by the authors after the above-mentioned theoretical foundation had been formulated. The experts were asked to assess each hypothesis individually and were given the opportunity to leave open comments to elaborate on their judgments or suggest improvements. The comments made by the experts in the first round proved to be highly constructive. Based on their input, the wording of several hypotheses was revised, one was removed, and two new ones were added. Despite only reporting the results of the second round here, we would like to provide a snapshot of the topics and nature of the comments in the first round: An example of the debate surrounding the precise choice of words can be seen in comment E1-5 on H1: “I find the word ‘increasingly’ difficult. I also find ‘hybrid production processes’ difficult. Both terms are open to individual interpretation, meaning that participants’ answers may be given from different perspectives. ‘Increasingly’ can mean ‘more

than today', for example, but also 'increasing per year'. When I heard 'hybrid production processes', I immediately thought of industrial car manufacturing;-)". The comment by E1-3 on H1 shows that virtual field trips can be seen in a very heterogeneous field of application: "I could definitely imagine this in the context of vocational training, but I see the challenge of costly infrastructure when it comes to expanding VFTs to informal learning in the private sphere." The difficulty of assessing future trends and identifying a vision for the future is also reflected in many of the comments. In a comment on H2, E1-9 criticises the slow pace of adaptation of the technical innovations that are already available: "I do expect this development, but not within this timeframe. Technological advances have not yet been adopted by manufacturers to the extent that would be necessary." E1-7 takes a similar view with regard to H3: "I don't believe that screens will have been replaced as the primary technology (overall) in 10 years' time, but I do think it's entirely possible that mobile XR devices will be so widely available by then that they can be used as the primary reception technology for VFTs." Comparisons of technical developments, as E1-2 does with regard to H4, are rather rare: "I think that AI-powered avatars will play a role, but I'm not sure that they will play a central role over other forms of personal interaction. It would also be useful to define personal interaction." Finally, E1-5 offers a kind of normative grounding for the visions provided in the hypothesis in relation to H4: "Largely replaced' means that more than 50 % of screen formats will be in XR. I have been researching XR for over 20 years and since I began my research, it has always been said that 'in 10–20 years, XR will replace 2D formats'. That has not happened and, in my opinion, will not happen easily – the effort required for a 2D application is SIGNIFICANTLY less than for an XR application, so there will always be 2D applications – more than 50 %."

Building on the qualitative feedback from the first round, the revised set of 14 hypotheses was evaluated in a quantitative second round via the online survey tool SoSci. The survey was distributed via email to 79 experts, and a total of 24 valid responses were received. Each hypothesis was rated on a 5-point Likert scale (1 = completely agree to 5 = completely disagree). The aim of this round was to identify consensual trends regarding the future development of VFTs. Table 2 presents the hypotheses along with their corresponding levels of agreement, ordered by mean value ($N = 24$). We collected data on the demographics of the experts invited to the second round based on their discipline, position, and country by analysing their professional web presences (see Appendix A). The

disciplines of educational science (35 %), educational technology, XR (extended reality) (23 %), and computer science (11 %) achieved double-digit relative frequencies. The positions of professor (56 %), doctoral student (15 %), and researcher (14 %) were the most frequently represented. Among the countries in which the invitees worked, Germany (41 %), the United Kingdom (13 %) and the United States (11 %) achieved double-digit relative frequencies. As the responses were anonymised, the demographics of the participating experts could not be determined and must be inferred from those of the invited experts.

The results indicate a clear tendency towards technological and didactical innovation in VFTs, especially regarding media integration and adaptive learning. Organizational developments concerning workflows were assessed positively, whereas deeper structural changes were met with noticeable scepticism. Estimations for all hypotheses besides H3, H13 and H14 differed significantly from the scale mean.

Technical Prospects. The strongest agreement was recorded for H1 – *Design and Production* ($M = 1.4$), underlining that VFTs will increasingly integrate various media technologies – such as 360° imagery, CGI, laser scans, and AI-generated content. This reflects current trends in immersive media convergence. Similarly, H4 – *Interactive Avatars* ($M = 2.0$) and H5 – *Physical Data* ($M = 2.2$) received positive ratings, pointing to anticipated improvements in realism and personalization, driven by big data and AI technologies. In contrast, H3 – *Reception stating* mobile XR as the main access point ($M = 2.7$) received the least agreement among the technical hypotheses, reflecting ongoing concerns about the breakthrough and broader adoption of mobile XR.

Didactical Prospects. High consensus was also evident for H9 – *Learning Analytics* ($M = 1.8$), emphasizing the relevance of adaptive, data-driven learning experiences. H10 – *Adaptive Visual Representations* ($M = 1.9$) further confirms expert confidence in the didactic value of spatial understanding and fluid perspective shifts in virtual environments. While H12 – *Collaborative Learning* ($M = 2.3$) was moderately supported, H14 – *Accessibility* ($M = 2.8$) showed the lowest agreement overall, indicating that the vision of fully inclusive immersive learning environments is still seen as difficult to realise.

Organizational Prospects. Strong support for H8 – *Development Tools* ($M = 1.9$) reflects optimism that AI-based, user-friendly tools will enable broader participation in VFT creation, even by non-experts. Similarly, H2 – *Content Preparation* ($M = 2.1$) points to a growing expectation that modular integration of pre-produced elements will become the norm, easing development efforts. By contrast,

Table 2: Hypotheses and their consent on 5-point Likert scales (1: i completely agree – 5: i completely disagree) ordered by mean value (N=24).

Hypotheses <i>Technical Prospects (TP), Didactical Prospects (DP), Organisational Prospects (OP)</i>	M (SD)	Sig.	Median
H1 – Design and Production (TP) <i>By 2035, Virtual Field Trips (VFTs) will be implemented through the combination of various technologies – such as 360° media, photogrammetry, CGI (Computer Generated Imagery), laser scans, and AI-generated content – to make real-world locations accessible in a cross-media and highly detailed way.</i>	1.4 (0.56)	M*** L***	1
H9 – Learning Analytics (DP) <i>By 2035, learning analytics systems will be specifically used in VFTs to individually evaluate learning progress and provide personalized content based on that analysis.</i>	1.8 (0.96)	M*** F* L***	2
H8 – Development Tools (OP) <i>By 2035, scalable, user-friendly (and AI-supported) tools will have simplified the development of VFTs to such an extent that even individuals with limited technical knowledge will be able to independently create learning environments.</i>	1.9 (0.70)	M*** F** L**	2
H10 – Adaptive Visual Representations (DP) <i>By 2035, technological advancements – such as high-resolution visualizations – will enable VFTs to support fluid perspective shifts between micro and macro levels (e.g., close-up views and overviews).</i>	1.9 (0.83)	M*** F* L**	2
H4 – Interactive Avatars (TP) <i>By the year 2035, AI-driven avatars will play a central role in supporting personal interaction within VFT-based learning processes.</i>	2.0 (0.93)	M*** F** L*	2
H2 – Content Preparation (OP) <i>By 2035, software solutions for designing VFTs will have evolved to the point that preproduced content can be integrated – self-created content will no longer be strictly necessary.</i>	2.1 (0.97)	M*** F** L*	2
H11 – Recommender-Based Experiences (OP) <i>By 2035, VFT platforms will feature recommender systems that suggest suitable VFTs based on individual interests, learning progress, and prior thematic knowledge.</i>	2.1 (0.95)	M*** F** L*	2
H5 – Physical Data (TP) <i>By the year 2035, VFTs will integrate physical (real-time) data to represent virtual scenarios in a more realistic way. This will help to better illustrate the underlying processes.</i>	2.2 (0.85)	M*** F*** L**	2
H6 – Immersive Live Streaming (TP) <i>Live streaming will become a standard tool for VFTs by 2035: learners will be able to “participate” at remote locations in real time and communicate live with on-site participants via 360° video and VR interaction, enabling, for example, virtual school trips with real-time experiences.</i>	2.3 (0.94)	M** F***	2
H12 – Collaborative Learning (DP) <i>By 2035, collaborative learning in VFTs (e.g., via social VR platforms) will be a standard part of the didactic repertoire.</i>	2.3 (1.05)	M** F***	2
H7 – Delivery (OP) <i>By 2035, central platforms – modeled on MOOC systems (Massive Open Online Courses) – will enable the broad provision and systematic use of modularly structured VFTs.</i>	2.4 (0.95)	M** F***	2
H13 – Integration into Teacher Education (OP) <i>By 2035, the creation and didactic use of VFTs will be an established part of teacher training and professional development for educators and teaching professionals.</i>	2.6 (1.03)	F***	3
H3 – Reception (TP) <i>By 2035, VFTs will predominantly – i.e., in more than 50 % of cases – be flexibly accessed via mobile XR headsets, which mediate situationally between physical and virtual environments and will have replaced traditional screen formats as the primary reception technology.</i>	2.7 (1.03)	F***	3
H14 – Accessibility (DP) <i>By 2035, VFTs will, through inclusive design, enable all learners – regardless of physical, cognitive, or social conditions – to benefit from full and equal participation.</i>	2.8 (1.14)	F***	3

Sig.: T-test against the scale mean (3): M**p <0.01; M***p <0.001. Wilcoxon signed ranks test against hypothesis 1 (design and production): F* p <0.05; F**p <0.01; F***p <0.001. Wilcoxon signed ranks test against hypothesis 14 (accessibility): L* p <0.05; L**p <0.01; L***p <0.001. The raw data can be obtained on request from the corresponding author.

H7 – *Delivery via Central Platforms* ($M = 2.4$) and H13 – *Integration into Teacher Education* ($M = 2.6$) showed lower consensus, suggesting that systemic changes within educational institutions are seen as more dependent on external factors such as policy, infrastructure, and professional training measures.

5 Conclusion, limitations and outlook

In addressing the research questions, this study shows that the future development of VFTs is closely linked to ongoing technological innovation. The concept of affordances provides a useful lens to structure this relationship across the dimensions of technology, experience, and learning. The Delphi findings suggest that VFTs are likely to become even more immersive, adaptive, and learner-centred – driven by advances in integrated media technologies, user-friendly authoring tools and AI-based personalization. However, the study also reveals a tension between technological progress and institutional implementation. This pattern reflects what Ogburn³⁹ described already over a century ago as a “cultural lag,” a phenomenon whereby technological advancements outpace the ability of social and institutional systems to adapt. As Fisher and Wright⁴⁰ note, such lags remain a defining feature of contemporary innovation landscapes. In line with this, the Delphi results show strong consensus on technical developments (e.g., H1 – Design and Production; H8 – Development Tools; H9 – Learning Analytics), but significantly lower agreement on organizational readiness, such as the integration of VFTs into teacher training (H13) or the establishment of centralized delivery platforms (H7). This suggests that while the technological foundations for next-generation VFTs are rapidly evolving, their large-scale implementation in educational systems may be delayed by organizational constraints and slow pedagogical adaptation.

Limitations. Despite these insights, the study has several limitations. Delphi methods depend on expert judgments and are thus influenced by the perspectives, experiences, and disciplinary backgrounds of the selected sample. With 24 respondents, the study provides solid trends, but the generalizability may remain limited. Given the panel size ($n = 24$), the statistical power of the second-round tests is limited, especially for small effects; therefore, non-significant differences should not be interpreted as evidence of no effect. Although the Delphi design enabled structured consensus-building among experts on future

developments, the study did not include any follow-up discussions or focus group formats. Such qualitative exchange formats could have provided additional insights into the reasoning behind individual expert judgements. At the same time, focus groups can be seen as an optional complement rather than a core element of Delphi studies, which deliberately relies on iterative, partially anonymised individual judgements in order to avoid group dynamics and dominance effects.^{27,41} Future studies could nevertheless supplement Delphi surveys with interviews or group discussions in the spirit of a mixed-methods approach.

Moreover, predictions about the future are inherently uncertain and methodologically challenging, as they always depend on the respondents’ current perspective. Future projections are always, to some extent, a glance into the crystal ball and should not be understood as definitive predictions, but rather as structured approximations of possible futures, at the time of the study. Further, the divergences observed in the study with regard to agreement on individual hypotheses, such as H14, are to be expected and reasonable for the Delphi method. The aim of the Delphi method is not only to establish consensus, but also to highlight different, sometimes contradictory assessments of complex, future-oriented issues.^{27,41} This is also reflected in our expert pool, which covers a broad range of disciplines and institutional roles and includes experts working in multiple countries (Appendix A). Such divergences can be attributed, among other things, to the heterogeneity of the experts, particularly with regard to their professional backgrounds, institutional roles, and practical experience.^{42,43} Hypotheses with normative or structural implications – such as H14 on the issue of accessibility – are particularly susceptible to divergent assessments, as they imply not only technological feasibility but also institutional, organizational, and ethical prerequisites. In addition, the literature indicates that different understandings of terms within a panel can increase the variance of judgments, especially when hypotheses are formulated openly in order to reflect a spectrum of perspectives.⁴⁴ The observed differences in agreement should therefore be understood less as a methodological weakness and more as an indication of real tensions between technological potential, educational objectives, and organizational conditions, the resolution of which is crucial for further research and discussion.

Transferability. The results of our study appear to be neither limited to a particular level of education nor to a particular educational technology but can be interpreted at the level of fundamental technological, experiential, and learning-related affordances. The identified potential of VFTs – in particular exploratory, self-directed, and

situated learning, as well as the importance of “place” and “scale” – is relevant in principle at different levels of education, such as primary and secondary education, higher education, or continuing vocational education and training. At the same time, the study shows that the implementation and sustainable integration of VFTs is limited less by technological or didactic factors than by organizational constraints. Differences in curricular requirements, institutional autonomy, resource availability, and teacher education influence the extent to which VFTs may be used. The transferability of the results should thus be considered limited: while the underlying affordances may also be applied to other immersive educational technologies featuring a comparable exploratory and situational character, their institutional implementation remains highly context-dependent and requires appropriately adapted organizational and didactic concepts.

Future research should expand the empirical basis by incorporating the views of practitioners and learners. Mixed-method approaches, as well as longitudinal studies, could deepen our understanding of how VFTs are actually experienced and how they could develop in future. As immersive technologies and AI continue to evolve, research must critically examine not just what is technically feasible, but also what is pedagogically meaningful and institutionally viable. Nevertheless, VFTs are likely to remain a widely accessible learning tool with a reasonable production effort when compared to the learning outcomes that are achievable.

Acknowledgments: We thank the participating experts for their valuable feedback and ratings in the Delphi study.

Research ethics: Not applicable.

Informed consent: Informed consent was obtained from all individuals included in this study, or their legal guardians or wards.

Author contributions: V.E.: study conception and design (incl. the affordance-based theoretical framing), hypothesis development, data collection, analysis and interpretation of results, and writing – original draft. H.S.: study conception and design (historical development of VFTs), hypothesis development, data collection, analysis and interpretation of results, and writing – original draft. F.W.: figure design and preparation, hypothesis development, data collection, analysis and interpretation of results, and writing – review & editing. B.E.: statistical analysis, hypothesis development, data collection, analysis and interpretation of results, and writing – review & editing. M.F.: supervision, writing – review & editing. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Use of Large Language Models, AI and Machine Learning

Tools: The authors used DeepL, Grammarly, and ChatGPT to check grammar and improve readability; all content decisions remained with the authors.

Conflict of interest: The authors state no conflict of interest.

Research funding: None declared.

Data availability: The raw data can be obtained on request from the corresponding author.

Appendix A: Demography of experts

Round 2

Table A-1: Demography of the round 1 experts; multiple disciplines possible (ordered alphabetically).

Pseudonym	Discipline	Position	Country
E1-1	Criminology, Criminal Justice, and Legal Studies	Doctoral Student	USA
E1-2	Learning and Teaching with Media	Doctoral Student	Germany
E1-3	Computer Science, Education Science, Social Science	Doctoral Student	Germany
E1-4	Artificial Intelligence, Education Science, Innovation Management, Information Science	Doctoral Student	Germany
E1-5	Engineering Computer Science, XR	Professor	Germany
E1-6	Design, Media and Educational Science	Professor	Denmark
E1-7	Computer Science, Educational Technology	Researcher	Germany
E1-8	Education Science	Researcher	Germany
E1-9	Computer Science, Education Science, Educational Technology, XR	Professor	Germany

Round 2 Discipline

Table A-2: Disciplines of the invited round 2 experts; multiple disciplines possible, ordered by frequency de-scending.

Discipline	Freq. (Abs.)	Freq. (Rel.)
Education Science	28	35 %
Educational Technology	19	24 %
XR	18	23 %
Computer Science	9	11 %
Engineering Computer Science	7	9 %
Geography	7	9 %
Environmental Engineering	5	6 %
Civil engineering, HCI, Psychology	4 each	5 % each
Business Informatics, Communication Studies, Geology, Information Science, Physics	3 each	4 % each
Artificial Intelligence, Business Administration, Cognitive Science, Geoinformatics	2 each	3 % each
Architecture, Art, Astronomy, Digital Media, Innovation Management, Mathematics, Social Work, Urban Studies	1 each	1 % each

Position

Table A-3: Positions of the invited Round 2 experts, ordered by frequency descending.

Position	Freq. (Abs.)	Freq. (Rel.)
Professor	44	56 %
Doctoral Student	12	15 %
Researcher	11	14 %
Lecturer	6	8 %
Policy Advisor	3	4 %
Non-Academic Professional	2	3 %
Independent Researcher	1	1 %

Country

Table A-4: Countries of work of the invited Round 2 experts, ordered by frequency descending.

Country	Freq. (Abs.)	Freq. (Rel.)
Germany	32	41 %
United Kingdom	10	13 %
United States	9	11 %
Switzerland	5	6 %
Canada, Denmark	3	4 %
Estonia, Indonesia, Netherlands, New Zealand, Turkey	2	3 %
Australia, China, Czechia, Finland, Romania, Sweden, Taiwan	1	1 %

References

1. Knutson, K. Field Trips are Valuable Learning Experiences. *Informal Sci.* **2016** [Online]. <https://informalscience.org/field-trips-are-valuable-learning-experiences/>.
2. Bitgood, S. School Field Trips: An Overview. *Visitor Behav.* **1989**, *4* (2), 3–6.
3. Behrendt, M.; Franklin, T. A Review of Research on School Field Trips and their Value in Education. *Int. J. Environ. Sci. Edu.* **2014**, *9* (3), 235–245.
4. DeWitt, J.; Storksdieck, M. A Short Review of School Field Trips: Key Findings from the Past and Implications for the Future. *Visitor Studi.* **2008**, *11* (2), 181–197.
5. Woerner, J. J. Virtual field trips in the earth science classroom. In *Proceedings of the 1999 Annual International Conference of the Association for the Education of Teachers in Science*; Association for the Education of Teachers in Science (ASTE): Greenville, NC, 1999; pp. 1232–1244.
6. Tuthill, G.; Klemm, E. B. Virtual Field Trips: Alternatives to Actual Field Trips. *Int. J. Instr. Media* **2002**, *29* (4), 453.
7. Hurst, S. D. Use of ‘Virtual’ Field Trips in Teaching Introductory Geology. *Comput. Geosci.* **1998**, *24* (7), 653–658.
8. Spicer, J. I.; Stratford, J. Student Perceptions of a Virtual Field Trip to Replace a Real Field Trip. *J. Comput. Assist. Learn.* **2001**, *17* (4), 345–354.
9. Abdul Rahim, E. E. Visiting a Milk Factory Without Gumboots: Students’ Attitudes Towards a Virtual Field Trip. In *Proceedings of 2013 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALF)*; IEEE: Bali, Indonesia, 2013; pp. 423–428.
10. Ehlenz, M.; Maquet, F. 19 squared – Interaktive XR-Inhalte gemeinsam selbst erstellen. In *Proceedings of DELFI 2024*; Gesellschaft für Informatik e.V.: Bonn, Germany, 2024; pp. 533–537.
11. Pham, H. C. Virtual Field Trip for Mobile Construction Safety Education Using 360-degree Panoramic Virtual Reality. *Int. J. Eng. Edu.* **2018**, *34* (4), 1174–1191.
12. Pirker, J.; Dengel, A. The Potential of 360° Virtual Reality Videos and Real VR for Education – A Literature Review. *IEEE Comput. Graph. Appl.* **2021**, *41* (4), 76–89.
13. Prahara, G. A.; Rezaldi, M. Y. The Examining of 360° Media that Support Learning in Field of Education. *KnE Soc. Sci.* **2025**, *10* (3), 83–90.
14. Rosendahl, P.; Klein, M.; Wagner, I. Immersive Training for Movement Sequences: the Use of 360° Video Technology to Provide Poomsae Training in Taekwondo. *J. Phys. Edu. Sport* **2022**, *22* (10), 2318–2325.
15. Shadiev, R.; Yang, L.; Huang, Y. M. A Review of Research on 360-degree Video and its Applications to Education. *J. Res. Technol. Edu.* **2022**, *54* (5), 784–799.
16. Fink, M. C.; Eisenlauer, V.; Ertl, B. What Variables are Connected with System Usability and Satisfaction? Results from an Educational Virtual Reality Field Trip. *Comput. Edu.: X Real.* **2023**, *3*, 100043.
17. Wolf, M.; Wehking, F.; Söbke, H.; Montag, M.; Zander, S.; Springer, C. Virtualised Virtual Field Trips in Environmental Engineering Higher Education. *European J. Eng. Edu.* **2023**, *48* (6), 1312–1334.
18. Mulders, M. Learning About Victims of Holocaust in Virtual Reality: The Main, Mediating and Moderating Effects of Technology,

- Instructional Method, Flow, Presence, and Prior Knowledge. *MTI* **2023**, 7 (3), 28.
19. Taylor, A. H.; Wallgrün, J. O.; Knapp, E. E.; Klippel, A.; Sánchez, J. J. Being There: Effectiveness of a 360-degree Virtual Tour for Increasing Understanding of Forest Treatments for Fire Hazard Reduction in California, USA. *Fire Ecol* **2025**, 21 (1), 2.
 20. Eisenlauer, V.; Fink, M. C.; Hart, L.; Ertl, B. Enhancing Educational Engagement: Evaluating the Potential of 360° Virtual Field Trips in Immersive Learning Environments. In *Presented at the Proceedings of the 19th International Conference of the Learning Sciences-ICLS 2025*; International Society of the Learning Sciences: Helsinki, Finland, 2025; pp. 2985–2987.
 21. Averbeck, F.; Leifeling, S.; Müller, K.; Schoenfelder, T. Virtual Reality in Social Work Teaching — Two Approaches to 360° Videos and Collaborative Working. *EJEL* **2024**, 22 (3), 111–123.
 22. Eiris, R.; Wen, J.; Gheisari, M. iVisit-Collaborate: Collaborative Problem-Solving in Multiuser 360-degree Panoramic Site Visits. *Comput. Edu.* **2022**, 177, 104365.
 23. Montag, M.; Söbke, H.; Wolf, M.; Wehking, F.; Kraft, E. Room for Collaboration: Analyzing Group Learning in Spatial Digital Learning Environments. *Comput. Edu. Open* **2025**, 9, 100308.
 24. Wolf, M.; Hartwig, S.; Steinhöfel, G.; Söbke, H.; Kraft, E. Panoramaviewer — a Framework for Educational Collaborative Virtual Field Trips. In *2024 International Symposium on Multimedia (ISM)*; IEEE: Tokyo, Japan, 2024; pp. 269–274.
 25. Gibson, J. J. *The Ecological Approach to Visual Perception: Classic Edition*; Psychology press (Taylor & Francis Group): New York, 2014.
 26. Zhao, J.; Klippel, A. scale - Unexplored Opportunities for Immersive Technologies in Place-based Learning. In *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*; IEEE: Osaka, Japan, 2019; pp. 155–162.
 27. Häder, M. *Delphi-Befragungen: Ein Arbeitsbuch*, Vol. 3, korr. Aufl; Springer VS, 2014. [Online].
 28. Domagk, S.; Schwartz, R. N.; Plass, J. L. Interactivity in Multimedia Learning: An Integrated Model. *Comput. Hum. Behav.* **2010**, 26 (5), 1024–1033.
 29. Petersen, G. B.; Petkakis, G.; Makransky, G. A Study of How Immersion and Interactivity Drive VR Learning. *Comput. Edu.* **2022**, 179, 104429.
 30. Wolf, M.; Wehking, F.; Montag, M.; Söbke, H. 360-based Virtual Field Trips to Waterworks in Higher Education. *Computers* **2021**, 10 (9), 118.
 31. Klippel, A.; Zhao, J. J.; Oprean, D.; Wallgrün, J. O.; Stubbs, C.; La Femina, P. The Value of Being There: Toward a Science of Immersive Virtual Field Trips. *Virtual Real.* **2020**, 24, 753–770.
 32. Fink M. C.; Pankow N.; Watzka B.; Spallek M.; Ertl B.; Schaum C. Virtuell auf die Kläranlage mit erKlär-VR: Vorstellung einer VR-Lernumgebung und Erfahrungen aus dem Schuleinsatz. *KA Korrespondenz Abwasser, Abfall* **2024**, 71 (8), 643–650.
 33. Makransky, G.; Mayer, R. E. Benefits of Taking a Virtual Field Trip in Immersive Virtual Reality: Evidence for the Immersion Principle in Multimedia Learning. *Educ. Psychol. Rev.* **2022**, 34, 1771–1798.
 34. Rupp, M. A.; Kozachuk, J.; Michaelis, J. R.; Odette, K. L.; Smither, J. A.; McConnell, D. S. The Effects of Immersiveness and Future VR Expectations on Subjective Experiences During an Educational 360° Video. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2016**, 60 (1), 2108–2112.
 35. Renkl A.; Gruber H.; Mandl H. Situated Learning in Instructional Settings: From Euphoria to Feasibility. **1996**.
 36. Khan, H.; Lee, G. A.; Hoermann, S.; Clifford, R. M. S.; Billingham, M.; Lindeman, R. W. Evaluating the Effects of Hand-Gesture-Based Interaction with Virtual Content in a 360° Movie. In *Proceedings of the International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments (ICAT-EGVE 2017)*, The Eurographics Association, Adelaide, Australia, 2017, pp. 63–70.
 37. Fink, M. C.; Robinson, S. A.; Ertl, B. AI-Based Avatars are Changing the Way We Learn and Teach: Benefits and Challenges, Presented at the Frontiers in Education, Frontiers Media SA, **2024**, 9, 1416307; <https://doi.org/10.3389/feduc.2024.1416307>.
 38. Shuta, A.; Shouta, S.; Nguyen, D. Design and Implementation of a Web-Based Multi-User Virtual Reality System for Virtual Tour of Remote Places. In *Proceedings of the 15th International Workshop on Immersive Mixed and Virtual Environment Systems*; ACM: Vancouver BC Canada, 2023; pp. 34–36.
 39. Ogburn, W. F. *Social Change: With Respect to Culture and Original Nature*; Huebsch: New York, 1922.
 40. Fisher, D. R.; Wright, L. M. On Utopias and Dystopias: Toward an Understanding of the Discourse Surrounding the Internet. *J. Comput.-Mediated Commun.* **2001**, 6 (2), JCMC624.
 41. Linstone, H. A.; Turoff, M. *The Delphi Method*, Vol. 1975; Addison-Wesley: Reading, MA, 1975.
 42. Beiderbeck, D.; Frevel, N.; von der Gracht, H. A.; Schmidt, S. L.; Schweitzer, V. M. Preparing, Conducting, and Analyzing Delphi Surveys: Cross-Disciplinary Practices, New Directions, and Advancements. *MethodsX* **2021**, 8, 101401.
 43. Niederberger, M.; Spranger, J. Delphi Technique in Health Sciences: A Map. *Front. Public Health* **2020**, 8, 457.
 44. Khodyakov, D.; Grant, S.; Denger, B.; Kinnett, K.; Martin, A.; Peay, H. Practical Considerations in Using Online modified-Delphi Approaches to Engage Patients and Other Stakeholders in Clinical Practice Guideline Development. *The Patient-Patient-Cent. Outcomes Res.* **2020**, 13 (1), 11–21.

Supplementary Material: This article contains supplementary material (<https://doi.org/10.1515/icom-2026-0010>).