



Article UNBODY: A Poetry Escape Room in Augmented Reality

Fridolin Wild ¹,*¹, Lawrence Marshall ^{2,3}, Jay Bernard ^{2,3}, Eric White ^{2,3} and John Twycross ⁴

¹ Institute of Educational Technology, The Open University, Milton Keynes MK7 6AA, UK

² School of Engineering, Computing, and Mathematics, Oxford Brookes University, Oxford OX3 0BP, UK; lwk.marshall@gmail.com (L.M.); suresh@sprungsultan.com (J.B.); ewhite@brookes.ac.uk (E.W.)

- ³ School of English and Modern Languages, Oxford Brookes University, Oxford OX3 0BP, UK
- ⁴ UCL Knowledge Lab, Department of Culture, Communication and Media, UCL Institute of Education, University College London, London WC1N 3QS, UK; j.twycross@ucl.ac.uk
- * Correspondence: f.wild@open.ac.uk; Tel.: +44-(0)1908-858-885

Abstract: The integration of augmented reality (AR) technology into personal computing is happening fast, and augmented workplaces for professionals in areas such as Industry 4.0 or digital health can reasonably be expected to form liminal zones that push the boundary of what currently possible. The application potential in the creative industries, however, is vast and can target broad audiences, so with UNBODY, we set out to push boundaries of a different kind and depart from the graphic-centric worlds of AR to explore textual and aural dimensions of an extended reality, in which words haunt and re-create our physical selves. UNBODY is an AR installation for smart glasses that embeds poetry in the user's surroundings. The augmented experience turns reality into a medium where holographic texts and film clips spill from dayglow billboards and totems. In this paper, we develop a blueprint for an AR escape room dedicated to the spoken and written word, with its open source code facilitating uptake by others into existing or new AR escape rooms. We outline the user-centered process of designing, building, and evaluating UNBODY. More specifically, we deployed a system usability scale (SUS) and a spatial interaction evaluation (SPINE) in order to validate its wider applicability. In this paper, we also describe the composition and concept of the experience, identifying several components (trigger posters, posters with video overlay, word dropper totem, floating object gallery, and a user trail visualization) as part of our first version before evaluation. UNBODY provides a sense of situational awareness and immersivity from inside an escape room. The recorded average mean for the SUS was 59.7, slightly under the recommended 68 average but still above 'OK' in the zone of low marginal acceptable. The findings for the SPINE were moderately positive, with the highest scores for output modalities and navigation support. This indicated that the proposed components and escape room concept work. Based on these results, we improved the experience, adding, among others, an interactive word composer component. We conclude that a poetry escape room is possible, outline our co-creation process, and deliver an open source technical framework as a blueprint for adding enhanced support for the spoken and written word to existing or coming AR escape room experiences. In an outlook, we discuss additional insight on timing, alignment, and the right level of personalization.

Keywords: augmented reality; arts; poetry; escape room

1. Introduction

UNBODY is an augmented reality (AR) HoloLens installation co-created by a crossdisciplinary team spanning the arts, computer sciences, and humanities and featuring texts and images by Jay Bernard. The AR experience creates a holographic landscape that focuses on multisensory experiences of texts, and, as such, represents a major departure from most AR installations that prioritize visuals (see Supplementary video S1). By providing an open source blueprint for an AR poetry escape room released under an MIT license on GitHub, we enable others to benefit from our findings, adding a poetic perspective to existing AR



Citation: Wild, F.; Marshall, L.; Bernard, J.; White, E.; Twycross, J. UNBODY: A Poetry Escape Room in Augmented Reality. *Information* **2021**, *12*, 295. https://doi.org/10.3390/ info12080295

Academic Editors: Ramon Fabregat, Jorge Bacca-Acosta and N. D. Duque-Mendez

Received: 13 May 2021 Accepted: 15 July 2021 Published: 26 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). escape rooms or facilitating the development of new ones that are more focused on the spoken and written word and enriching the repertoire for AR escape rooms with several new forms and zones of interaction. We decided that our experience had to encourage user experimentation while creating a clearly defined path through the installation. The resulting application was implemented in Unity3D, using Microsoft's Mixed Reality ToolKit (MRTK) and Vuforia's Marker Tracking ToolKit. In this paper, we describe and document these novel methods and strategies with which we co-created UNBODY, and then we evaluate their level of usability and spatial experience with a system usability scale (SUS, [1]) and a spatial interaction evaluation (SPINE, [2]). In the discussion that follows, we show how escape room modelling, poetry game mechanics, and spatial audio enrich interactions in AR installations, and we also identify where further research is required.

The key technical frameworks for UNBODY are virtualities and, especially, escape rooms. In current usage, extended reality, or XR, refers to all categories of digitally simulated environments. Virtual reality (VR) is where the user immerses in a fully rendered environment simulation [3]. By contrast, AR refers to enhancing human perception with additional computer-generated sensorial input to create a new user experience [4]. From a more technical view, AR often refers to the strategic combination of real and digital virtual elements, providing responsive, real-time interaction and registering content in 3D space [5]. Mixed reality (MR) is not yet such an established term; it is often used synonymously with AR, sometimes more strongly connected to the use of spatial mapping or the alignment between a virtual overlay and a real environment [6]. Often, AR is working predominantly with the visual sense. In [7], however, the authors emphasized the importance of spatial audio for AR, concluding that "scaled 3D sound significantly increases the accuracy of depth judgments and shortens the searching task completion time" by a factor of around 2.6 times. Spatial audio is an important dimension of the UNBODY experience and part of our challenge to both AR and escape room paradigms.

Escape rooms and AR seem like a perfect match, warranting a myriad of potential applications in education, entertainment, and the creative industries. They provide a narrative scheme and physical/kinetic route through exhibition spaces and interactions with objects in those spaces. As of now, however, very limited research has been conducted on how to create such AR rooms. [8] surveyed 175 real-life escape room facilities around the globe. The author placed the first ever escape room in Kyoto/Japan in 2007. From there, he reported, the concept gained rapidly in popularity over the years 2012–2013, first in Asia. Then, via Europe and in particular Hungary, it spread to Australia, Canada, and the US. Earlier precursors included live-action role-playing, point-and-click adventure games, and treasure hunts [8]. He categorized escape rooms into mere collections of puzzles (13%), themed rooms (27%), puzzles that solely advance the narrative (21%), and (finally) puzzles as part of the narrative (39%). [9] then identified the key elements of escape rooms as finding clues, solving puzzles (or accomplishing other set tasks), and leaving the room before the timer runs out. Beyond that, only a few proposals that shed more light on the characteristics and design dimensions of escape rooms have been published. Escape room concepts are not the same as evacuation scenarios (see, e.g., [10]), though they certainly have a shared focus on moving users from one space to another, as well as similarities with the AR technologies (like markers). Escape rooms, however, typically do not have the sophisticated simulation engines and algorithms required, for example, for realistic flame visualization, smoke visualization, and flow prediction (ditto, p. 5).

Ref. [11] proposed a development framework for escape rooms for higher education, stressing that "themes and narratives set the game's atmosphere and lay the foundations of emotional investment and curiosity within the player." In their view, the theme impacts all other key characteristics of an escape room, namely puzzles, props, and design. They proposed a model for creating escape rooms for learning and behavior change along six design dimensions: participants, objectives, theme, puzzles, equipment, and evaluation. Nevertheless, their supporting evidence base was small, and further cases are required to study the framework's functionality and effectiveness. The framework puts emphasis

also on target audience selection aspects, learning objectives, and assessment ('evaluation') deliberations from an educational point of view, which are not relevant for our investigation. The additional differentiation along four themes and the four aspects of puzzle design, however, matter. The four listed themes are: escape mode ('leave the room'), mystery mode ("solve a mystery"), narrative design (interesting story), and stand-alone vs. nested ("one-off experience" or part of a series). Regarding puzzles, both their design and ability to reflect the learning objectives (which are not relevant here) matter, as does provided support with respect to instruction, manuals, and/or clues/hints.

The case study presented in [12] on two augmented reality escape rooms developed (Lost in Space; Escape the Cell) deconstructed educational escape games, proposing classification along the following design dimensions: number of participants, collaborative quest, age groups, level of difficulty, learning topic, instructor required, help availability, number of rooms, degree of virtuality, type of quests, time limit, and side quests. They identified a unique advantage of AR to work with complex 3D objects that are physically not available at scale or at such viewing angles. Moreover, with the help of virtuality, they determined that it is possible to simulate numerous rooms in a single physical space.

Ref. [13] assessed the benefits, characteristics, and performance measures of a teambuilding AR/MR escape room, AMELIO, arguing that the "merging or blending of social VR/MR technology and escape rooms has potential for professional practice and research of team performance." They proposed to break immersiveness down into four components: systems immersion, narrative immersion, spatial immersion, and social/empathic immersion (see Table 1). Systems immersion is related to the ability of the app to engage the user in the puzzles and mechanics of the game. Narrative immersion is about awakening a "compelling interest in the sequence of events." Spatial immersion refers to the ability to convey a sense of "presence in the location," while social and emphatic immersion helps create a "feeling of connectedness" to collaborators and game characters. In their view, each component equally contributes to establishing immersion. They added that effective audio is more important than visual aesthetics in this regard.

Table 1. Immersive dimensions [13].

Immersion	Meaning	
Systems Immersion	Engage with mechanics/puzzles of game	
Narrative Immersion	Interest in storyline and event sequence	
Spatial Immersion	Feeling present in the location of the game	
Social/Empathic Immersion	Feeling connected to other players and game characters	

Ref. [14] explored a situational AR escape room. They investigated collaborative experiences and studied the potential of escape rooms to build awareness, foster collaboration, and enhance communication, with great potential to shape leadership and group hierarchies, test conflict resolution, improve proxemics, and practice distributed cognition. Further, they identified basic levels of technology as "sensors, lights, or displays," adding that "as escape rooms evolve, it is likely that they will include an increasing amount of technology" and suggesting that future escape rooms likely will utilize more advanced equipment to create deeper levels of immersion. They foresaw in particular reactive escape room environments where the user picking up objects triggers audio-visual feedback cues. Thus, "players will be more likely to glance at the direction of the feedback and gain a greater sense of situational awareness." They also observed that "an object may reveal more information when held by a parent than her child," which allows the narrative to be broken down into parts that are linked and adapted to individuals and their characteristics. This process entails creating an escape room that virtually creates multiple characters per story. At the end of the experience, each player leaves the room with a similar story but adapted to a unique, individual interpretation. The application of more advanced puzzles or age-restricted content (e.g., to create a wider range of interest for families) could also be considered.

Ref. [15] explored duplicating an audience experience of an art performance using VR, assessing usability in a two-level framework by combining aspects of presence with quality evaluation typical for arts performances. They analyzed the level of presence (level 1), and, furthermore (level 2), engagement, information recall, and desire of viewing using four methods: a questionnaire instrument, interviews, EEG measurements, and head movement tracking. In their conclusion, VR was most favored for its sound and light, providing a better sense of presence and immersion. They found that VR participants were not always able to recite answers when questioned about specific elements of the storyline, concluding that aspects of the storyline may easily be missed in VR and require good highlighting and signposting to ensure they are not overlooked. They suggested introducing a certain level of orientation control to help viewers better follow plot.

Ref. [16] developed an interactive application for museum visitors using an AR view to create an immersive experience by superimposing generated 3D objects over real-life museum artworks, searching for related works of music, and using the generated 3D model as an audio visualizer. They used music to "make visitors more engaged without interfering with or being separate from their appreciation for traditional art forms".

Ref. [17] differentiated escape rooms in AR along the aspects of shop-system, puzzles, tangibles, and multiplayer. Shop-system refers to the ability to win points by solving puzzles that can be exchanged for additional hints in order to gain a competitive advantage. Puzzles refer to interactive (e.g., patterns visible only from specific angles through an AR portal or automata that must be turned into a specific state) and non-interactive puzzles (e.g., mathematical number sequences to be solved). Tangibles are their equivalent in the real world. For example, a pyramid puzzle with image targets on the faces of the 3D parts assembles into a recognizable marker. Among the multiuser experiences, they recommended differentiating between cooperative and competitive. While a hint-system and multiuser system were out of our scope, the storyline recommendations regarding puzzles and the environmental recommendations regarding the use of tangibles are very relevant to UNBODY.

Ref. [18] went further than the other presented AR escape rooms and integrated a client-based application plus server and further Raspberry Pi-based electronics to achieve a more natural interaction with objects in the real world via control through motors, e.g., the opening and closing of a leaves curtain and using deep learning with convolutional neural networks for improved computer vision to deliver cues automatically. They reported an accuracy rate of 94.9% when using 18,000 images in the training set (for 10 labels) and using 6000 images in the test dataset.

Ref. [19] described how designing an escape room with students forms the basis of a science-based curriculum at the American Museum of Natural History in New York. He emphasized the value of the concept of 'incomplete information,' which means that players do not have all information about the game's content or other players.

Ref. [20] developed an educational AR escape room to teach chemistry in playful way. In addition to the more active learning, they stressed the potential of adaptability for AR escape rooms regarding its complexity—turning the escape room from "discovery activity" to an experience at "specialized master level." Though it was deployed for a much different purpose, their work shows how AR escape rooms can be used to structure reticulated experiences with a variety of possible outcomes.

Ref. [21] ran a contest at the 3DUI conference for passive haptic feedback in AR escape rooms. The YouTube channel (https://www.youtube.com/playlist?list=PL5d4WtkvirVD4 vknKE3T0tf6BdHSO4H4A, accessed on 16 July 2021) provides insights into how the nine runners up used off-the-shelve and experimental trackers, haptic displays, and electronics to create intelligent props involved in game play, with focus on rotation, insertion, pressing, pulling/pushing, texture/material changes, joining objects, pressure/squeeze, and the feel characteristics of an object (shape, temperature, etc.). [22] described their contribution to the competition in more detail, documenting, how they designed and 3D-printed the developed props (e.g., blow torch, a 3D-printed valve, using HTC Vive trackers, and Arduino).

In a similar fashion, [23] proposed the use of electro-muscular stimulation (EMS) for applications in augmented reality escape rooms, exemplified by the application of an EMSenhanced manipulation of goose-neck lamps that symbolize levers operating a door that can only be forced into specific positional combinations using modulated muscle detents.

Ref. [24] investigated the use of Bluetooth low-energy beacons, stressing the importance of using spatial skills in the designed user experience.

Ref. [25] described a complex installation that uses costumes with RFID tags, a Kinect camera system, Arduinos, LED strips, a controlling computer, and smart phone apps. They warned that with AR installations, it is possible that "the novelty of the experience will override any other pleasures (or problems) with the design," emphasizing the importance of the learnability, i.e., allowing the user to progress easily from being unfamiliar, to being comfortable in use, to being an expert. They also indicated that recognition comes before affiliation, allegiance, and (finally) identification, indicating the greater challenge in creating cult-status around a game with "depth of play and long-term engagement in mind".

We can therefore summarize the state of the art as being along the four aspects of storyline, environment, visuals, and audio in order to guide the development of our artistic augmented reality poetry escape room (see the summary in Table 2). In the UNBODY cocreative process, we adapted the escape room format for the project's needs as it developed, and we targeted 'storyline' and 'audio' as our key areas for development and innovation. This was partly because some of Bernard's poetic texts were not entirely linear and partly (as with [20]) because the visual imagery invited experimentation and offered multiple possible outcomes.

Dimension	Key Recommendations	References
Storyline	Use systems, narrative, and spatial immersion to build up	[12]
	social/emphatic connectedness	[13]
	No time limit needed	[13]
	Define narrative early in design and development process	[11]
	Theme impacts on puzzles, props, and design; use theme to enhances the objectives of the game	[11]
	Puzzles can be interactive or non-interactive.	[17]
	Principle of 'incomplete information' facilitates development of inquiry skills.	[19]
	Adapt to user characteristics (like age) or mastery level	[14,20]
Environment	Vary teams-based familiarity (pre-existing vs. newly composed), vary roles or level of difficulty	[13]
	Headsets risk users missing key elements of the storyline	[15]
	Simulate several rooms in single physical space	[12]
	Progression should be apparent at all times throughout the experience	[11]
	Involvement of tangible objects is possible to relate the digital to the real-world space.	[17]
	Electronics can be used to modify the environment or create passive haptics.	[18,21]
	Provide support in form of instruction/manuals	[11]
	Use spatial skills of the user to identify game locations	[24]
	Establish a framework to use to design and develop upon	[11]
Visuals	Visual cues improve visual search performance	[7]
	Provide clues/hints	[11]
	Unique ability to use complex 3D objects physically not available at needed scale/viewing angle	[12]
	Picking up objects triggers audio-visual feedback cues	[14]
	Lighting effects strengthen immersion	[15]
Audio	Spatial audio cues significantly improve visual search performance	[7]
	Audio is more important to a user than visual aesthetics	[13]
	Sound effects to strengthen immersion	[15]
	Relate music works to visual objects	[16]

Table 2. Key recommendations for escape rooms.

Much is known in isolation about escape rooms, mixed/augmented/virtual reality, and even mixed/augmented/virtual escape rooms for education, but the combination of an arts performance and poetry game mechanics into an escape room in augmented reality remain underexplored. Moreover, besides that of [12], there are no escape rooms on head-mounted see-through displays (of the grade of Microsoft HoloLens). The text-based nature of UNBODY sets unique design challenges, solutions for which we describe in this article. Furthermore, our UNBODY ambition provides a unique opportunity to extend the work done so far with regarding the use of spatial audio in AR escape rooms.

2. Concept

The design concept for UNBODY applies escape room dimensions (see Table 2) to a poetic creation. Specifically, UNBODY investigates the ambiguity of our digital worlds and our physical bodies through AR texts. Bernard's work explores being black, queer, and trans in Britain [26–28]. In UNBODY, they use texts to break down users' conceptions of the boundaries that separate categories of waking and dreaming, physicality and digitality, and body and identity. In UNBODY, Bernard writes "we acknowledge possibility, not probability. This is the same world, but you are sleeping. Your living body relates to the world in a way that is unknowable to the conscious mind. Holding your sleeping self in your hand, this is how you move through a dream." In other words, UNBODY reconstructs a fictional dream state that can change as the user moves through the installation space, which is subdivided into three zones.

The authors of [11] stressed the importance of defining narrative early in the design and development process. Because the UNBODY text does not follow a clear-cut narrative progression, we needed to define 'narrative' in other ways, particularly using theme to enhance the objectives [11]. Physical space and objects were especially important factors when designing the app for a head-mounted display.

Our initial plans for UNBODY responded the concept presented in [17] that tasks and objectives can be interactive or non-interactive. The concept also evolved to accommodate logistical arrangements, which determined how we structured the sequence of experiences.

The size of the exhibition room was paramount to many requirements of the project, in particular for the HoloLens' tracking capabilities, layout for props, people capacity, and segmentation into parts of the experience. The choice of room seemed suitable for the application by being relatively large, thus allowing room for error and the modification of layout choices by using dividers and large panels to section off unwanted space and clad the big glass front.

The layout of the 'naked' room lacked in several areas, mainly and firstly that there was no natural flow for the user experience. Users would have to know where to stand and pause, as well as which sections to go to next. Secondly, the room layout offered little to no diversity for demonstrating extended functional AR capabilities. Finally, the provided space was longer than expected, risking issues with the spatial mapping capabilities (see Figure 1).



Figure 1. The exhibition room before walls and furniture were rearranged.

We planned for three zones created with room dividers. These zones were arranged in a 'tube' in which the experience takes place, framed by an entry and an exit tunnel, by means of which the audience can stream in and out. Figure 2 depicts the second design scribble of a master plan for the experience. Visitors pick up their headset from the table in the top left corner (see Figure 2). They immediately can see the Zone 1 trigger poster, which invites them to come closer. Activated by proximity, Zone 1 unveils and six posters with black and white images of the poet become reactive through trained Vuforia image targets. Whenever the user sees a target poster, a video overlay is automatically revealed on the poster. At the end of the tunnel of Zone 1, the trigger poster for Zone 2 displays a virtual trigger object guiding the user over to Zone 2. In Zone 2, the user picks up a 'totem,' an inanimate object, which comes to life when the user gazes upon it. The totem object, a cube, starts dropping the words of a poem, which is read in parallel into the ear of the spectator using spatial audio. Around the corner awaits Poster 3, which launches a virtual exhibition in Zone 3. Floating objects (spheres and tilt brush sculptures) appear, each of them fitted with a range-based audio activation that allows for the compilation, out of pre-recorded sentences, of a novel poem, engaging the user in the act of compiling it. When exiting, the user can turn back, and a final poster triggers the unveiling of the poem's pathway (the path the user has chosen to run into the virtual exhibition objects). In this manner, the user seeks the next trigger to activate the next text, audio event, film, or image. In this respect, "moving through the dream" in a free yet structured way becomes the narrative objective, rather than 'escaping' from an imagined threat or solving a particular 'puzzle' or set task; see also Table 3.



Figure 2. Design mock-ups: improved master plan.

Table 3. Original	l room pl	lan with zones.
-------------------	-----------	-----------------

Zone	Description
Enter	This is where users pick up their HoloLens and don it. Trigger posters mark the entry of each zone; on collision with an invisible collider around the AR camera (i.e., the head-set position in the room), these trigger posters switch to the relevant zone they sign post.
Zone 1	Video overlay posters: six posters positioned on the wall are trackable using the Vuforia engine in unity. Videos are superimposed as flat game objects on these posters, with the physical posters being placed in a defined order along the walls. Overlay objects are placed relatively close to their physical counterpart. Their absolute position in the world space is maintained using anchors in the spatial map.
Zone 2	A word dropper totem is augmented through the HoloLens and can be used to create words from the poetry that fall out, rendering them in the world space. Users pick up one of five cubes. Once visible by the camera, the word dropper 'box' begins to generate words. The words change over time to form a poem that falls out of the box into the world space.
Zone 3	There is a floating object gallery (orbs and 3D sculptures) within a random space of the zone. When the virtual exhibition objects are walked into, the containing content inside them can be viewed, and a separate audio narration automatically plays for each object through the spatial audio speakers of the HoloLens.
Exit	User trail visualization: a line is generated between objects within Zone 3, which shows the user the sequence in which they visited each floating object, visualizing the user's pathway in 3D space. The user can go back around Zone 3, triggering the content again if they choose.

"What if I said someone who loves you very much is watching over you? What if I said someone with your best interests at heart is listening to you breathe? The world is going on without you-the version of yourself you know the best. But here's a whole world that can't continue without you.

When you look at the world, with its stark black and white division, don't you wonder about the people who lived before, don't you wonder about the ones for whom dreaming, and waking were on the same continuum?" (excerpt from Bernard's poem).

3. Implementation

The UNBODY application was developed in Unity3D (2017.4.35f1 LTS and Windows SDK 10.0.18362), implemented in the C# programming language using Microsoft's Mixed Reality ToolKit (MRTK, 2017.4.3.0), as well as Vuforia's Marker Tracking SDK. As development for the HoloLens is still very sensitive to version numbers and their dependencies, we chose a long-term support version of Unity and stable releases of MRTK 1 'HoloToolkit' for development.

The application was released as open source under an MIT license, available free of charge at this link: https://github.com/fwild/PoetryEscapeRoom (accessed on 16 July 2021).

The application is under active development, so version numbers and dependencies are likely to change. Any changes in requirements will be updated accordingly and are listed in the README file in root directory of the GitHub repository.

The basic application logic is that each zone has a trigger poster (pose estimation established using Vuforia image target tracking), which superimposes a sphere object (with collider). The AR camera of the MRTK has been equally fitted with a collider, so that when the user walks the camera's spherical collider 'shield' into the poster overlay 3D object, the trigger fires and a *SceneManager* behavior script 'CollisionHandler.cs' is called to activate the new visuals or according spatial audio (and, if needed, deactivate 3D and audio content from the zone that the user departs).

Figure 3 depicts the UML sequence diagram for Zone 1's video overlays. The user approaches a poster and verifies whether Vuforia recognizes an image target. If an image is found, the *DefaultTrackableEventHandler* fires and the video is displayed and played. Whenever the tracking is lost, the video disappears and pauses. Vuforia uses extended tracking, utilizing the HoloLens's spatial mapping to stabilize tracking.

Figure 4 depicts a video playing over a (black and white) poster. Figure 5 shows the totem. Figure 6 depicts mixed reality capture of the word dropper from a third person perspective. Figure 7 shows the connecting lines the user sees when looking back from the exit onto Zone 3, visualizing the order in which the user has 'released' the poem's audios from the interactive tilt brush artworks.

The word dropper functionality (script 'Gnome.cs') uses a timer in the update loop to simultaneously drop both a word object (visually representing the word in the line of the poem being read) and play its audio recording. The timer checks word after word against its according timecode to then uses a ray cast from the position of the smart glasses to identify where the user is currently gazing. If the ray cast hits the spatial map from the depth camera room scan, the visual word game object is placed away from the hit point along the gaze direction, 10 cm closer to the user. If the ray cast hits another object (usually another word), then it places the word at the position of the totem plus 5 cm upwards and 1 m forward. The same thing happens if the ray cast hits nothing, which is typically the case when the spatial map has not been scanned because, e.g., the walls are too far away.



Figure 3. Sequence diagram for the image targets and video augmentations.



Figure 4. Video poster (Zone 1).



Figure 5. Totem (Zone 2).



Figure 6. Word dropper (Zone 2).



Figure 7. Path and tilt brush artwork (Zone 3).

The line path depicted in Figure 7 gathers its data in the same trigger script that plays the audio upon collision with the AR camera ('BallBehaviour.cs'). It uses a *LinePathManager.visitAsset()* method to log the indoor 3D position, which is then later visualized from the exit poster trigger by calling *LinePathManager.drawLines()* and a *LineRenderer* component.

4. Evaluation

To evaluate the system, we asked visitors to the exhibition to participate on a voluntary basis in an experiment by providing feedback in the form of a questionnaire filled out after to their experience of the exhibition. Since no personal data were collected, ethics clearance was not required. We used the System Usability Scale (SUS, see [1,29]) as well as its extension to augmented and mixed reality—the Spatial Interaction Evaluation metric (SPINE, [2])—to shed light on the usability and user experience of UNBODY and the viability of its approach. We used the Microsoft Hololens, alternating several devices during the exhibition to ensure they recharge while not in use. The devices were purchased from Microsoft in the UK.

For SUS, ref. [30] recommended an absolute bare minimum number of five users, which already yields a score within 6 points of the actual value 50% of the time (95% of the time with 17 points difference). However, he also recommended a sample size of 10 as a more realistic minimum sample size [31,32] in order to record reliable results. This is in line with recommendation of online calculators like the Usability Sample Size Calculator at BlinkUX.com [33], which recommends 10 while expecting 8 to show. With the same mathematical justification, we set the sample size recommendation in analogy for the SPINE. This included seven participants for the spatial interaction evaluation data and nine participants for the system usability scale data.

Figures 8 and 9 give an impression of the exhibition. We first gave all users a short introduction to the app, then helped them don the smart glasses (Microsoft HoloLens), and sent them off through the experience. Stewards were positioned near the trigger posters for Zones 2 and 3 to help guide the spectators onwards in case they seemed stuck. We did not require anyone to perform the gesture training, as all interaction with UNBODY is purely gaze based. The helpers in the middle of the experience instructed the user to pick up the word dropper cube.



Figure 8. The poet in the exhibition.



Figure 9. Bernard in the Zone 2 poem.

A system usability scale (SUS, see [1,29]) and a spatial interaction evaluation (SPINE, see [2]) were deployed to measure the usability of the system. A SUS breaks usability down into ten items, and a SPINE adds augmented reality-specific elements about the interaction and determining how well the implementation of the system supports control (SC), navigation (NAV), manipulation (MP), and selection (SEL) interaction tasks, as well as checking how input modalities (IM) and output modalities (OM) are supported. Both methods were used with a five-point Likert scale. For the SPINE, data coding converted the values to -2 (strongly disagree) and +2 (strongly agree). The middle, 0, indicated 'neither.' The SUS was mapped from 0 to 4, following the instruction of Sauro (2011). Negative formulated items were reversed in polarity during data coding (items 8, 11–13, 18 for

SPINE; all even items for SUS). The total SUS score was normalized to 100 by multiplying the mean with 2.5, as dictated by [30].

The selection of participants for both questionnaires was random and on voluntary basis. Participants could choose from one of two questionnaires, thus relaxing the time constraints needed to keep participants moving through the experience at a set pace. This included seven participants for the spatial interaction evaluation data and nine participants for the system usability scale data.

Regarding usability, the results of both the SUS and SPINE showed that the system was usable and the experience worked well for most of the participants.

The group of input modalities questions were judged with low values in the SPINE, which was not surprising because these item groups contain questions about the additional modalities that were not used by the app (voice, gestures, and controller). The most relevant items, 'Using gaze to trigger system functionality was difficult.' and 'I was able to understand what could be gaze-activated.', were rated with means of 0.29 and 0.57 respectively, at standard deviations of 1.60 and 0.98, respectively, so the suspected 'true' value for this item group was rather closer to the other constructs. The overall mean was 0.56 at a standard deviation of 1.08 with a median of 1; see Figure 10. The findings for the SPINE were overall moderately positive, with the highest scores given for output modalities and navigation support.



Figure 10. Results for the SPINE (negatively formulated items 8/11/12/13/18 reversed; SC = system control; NAV = navigation; MP = manipulation; SEL = selection; IM = input modalities; OM = output modalities; *n* = 7).

The recorded average mean for the SUS was 59.7, slightly under the recommended 68 average but still above 'OK' and therefore of low marginal acceptability. This shows room for improvement of the experience. For the reverse scaled item four in particular ('I think that I would need the support of a technical person to be able to use this system'), UNBODY did not perform very well (see Figure 11), resulting in a mean of two. In our view, the most likely interpretation of this finding for item four is that participants had no prior exposure to this new type of device (Microsoft HoloLens), thus not benefitting from any past experience in operating hands-free smart glasses with hand gestures and eye gaze. At the time of writing, HoloLens is still only available for retail for enterprise and developers. It is not available to be purchased by consumers on the open market, so the number of users in Oxford who had access to such device (if anyone at all) before is likely to be very limited.





The tested version also had no embedded guidance other than the interesting looking trigger orbs overlaid at an approximate 1.5 m distance from the zone trigger posters. Similarly, the lower average score on item three ('I thought the system was easy to use') and the reverse-scaled item ten ('I needed to learn a lot of things before I could get going with this system') indicate that some of the participants experienced these as issues, finding that it was not so easy to use and that they would need to learn more about the system functionality in order to get to satisfying use.

Regarding the performance, all aspects of design functionality worked well when demonstrated by the participants, but the experience could have been improved through clearer instruction on how to navigate the room directly embedded in the app. Participants could effectively interact with the environment and the application was responsive in interactions with participants. In our view, the lack of in-app instruction, in particular, and guidance can be improved. Moreover, the selection and input modality of Zone 3 can be significantly improved.

It was clear that many participants enjoyed the experience, the playfulness of the word dropper, the poetry whispered into their ears, and the video visuals, and they gained a broader perspective of the prospects AR technology has to offer. The posters, visual props, and ambient soundtrack composed by Bernard also helped create an immersive and theatrical experience that enhanced the emotional response of the participants.

To summarize, the SUS and SPINE results showed that the concept and components work but indicate room for improvement. Subsequently, we describe how we further developed the prototype into a second release.

5. Further Development of the Prototype

Following the exhibition and evaluation, we further improved the prototype to a second version. We used two of the tilt brush artworks to replace the abstract activation spheres of the zones' trigger posters (Figures 12 and 13), also scheduling the new Zone 3 with the word composer to start after the word band for Zone 2 spools out of the totem. The biggest shortcoming was the lack of in-app guidance and the difficulty in navigating Zone 3 with the 3D artwork. Therefore, we added in-app instruction and provided a floating box with text. Furthermore, we removed the tilt brush artwork (and connected 'chopped up' poems) to make room for a new experience, the word composer (Figures 14 and 15). We optimized Zone 2's timings, aligning words to millisecond precision of when the words are visually dropped. Moreover, we modified the drop locations so that the user's gaze cursor now allows words to be placed on walls, if the user's gaze is focused on the distance rather than on the totem box. We also optimized the physics of the word dropper.

The new word composer offers the user 27 possible combinations of prefix (trans-, pre-, and un-), stem (body, conscious, and dream), and suffix (-ing, -ectomy, and -ly). For all possible 27 combinations, one definition was written by the poet and added to the system to be displayed upon completion of a compound by the user. Relevant word groups are highlighted in red and show a gaze-activation progress bar to indicate that selection is happening. At the end, the instruction box asks the user to make their way out and return the headset.



Figure 12. New trigger object (Zone 1).



Figure 13. New trigger object (Zone 2).



Figure 14. Word composer (Zone 3).



Figure 15. Definition provided for the new compound un-body-ily (Zone 3).

6. Summary and Conclusions

UNBODY is a blueprint for a poetry-focused escape room experience with AR, originally written for the Microsoft HoloLens. It is the first of its kind. The escapED framework was useful to help focus the initial development of UNBODY, turning attention to theme, puzzles, and equipment. [11] proposed to distinguish themes into four types: escape mode, mystery mode, narrative design, stand-alone/nested. UNBODY is a mix of 'narrative design' and 'stand-alone/nested.' Beyond these, UNBODY is exploratory and a ludic experimentation with text, testing out the possibilities of language in an extended reality.

UNBODY is the result of an interdisciplinary collaboration built with co-creative practices. The methodology we deployed in UNBODY adapts four classic processes established by the Italian collective Artway of thinking and outlined in [34]: analysis (observation), concept generation (co-generation), restitution (action on feedback), and integration (metabolization of the innovation). We repeated these steps at key development stages to democratically incorporate each voice into the final production, alongside end-users' creative input and feedback, according to what was technically and logistically possible. One upshot of this approach was that we quickly began to stretch standard understandings of escape room conventions. We had to 'escape the escape room,' particularly around standard notions of narrative.

Narratives serve to maintain the user's interest in the experience. However, we found that this interest can take multiple forms and be guided just as much by exploration and searching for the next experience as by problem solving, characterization, or other standard escape room scenarios. It was also possible to enhance user immersion in the experience by means of multimedia framing.

In agreement with the findings in [7,13,16], audio quality has a significant impact on the level of immersion achieved. Though effective audio satisfies and involves the standard user, those who suffer from hearing impairments will likely be negatively impacted. Future work could elaborate recommendations on how to achieve a good balance between features and the according accessibility. As the experience focuses on storyline, the stand-alone design makes sense, as the experience is played by a single user in a single-purpose environment that contains a single state of progression.

Puzzles (or puzzle-like set pieces such as the word box and the word composer) complement the theme of poetry by attracting the attention and interest of the users and creating memorable experiences. Each puzzle should provide clear instructions or follow natural rules for engagement so that the puzzle is easy to use and understand. The quality of the interaction should complement the general themes of the experience.

The final area of the framework is equipment, which considers environment, location, and equipment requirements. Performance becomes key when lighting, props, space, actors, and environment items are chosen to affect the experience.

This is in line with findings of other related studies. In [35], for example, proposed deploying user-centered design when developing for heterogeneous audiences, using loops and narratives as methods of audience engagement, leveraging interdisciplinarity to bring together the best practices in museum pedagogy and creative industry. From a user experience perspective, we can conclude that timing and alignment are essential for such a poetic, word-focused experience. For example, optimizing the word drop times required several iterations in development, focusing on aligning the visual correlates with the audio experience, until the user experience was right. However, this is not only about the precise timings of delivery in different, parallel modalities.

This is also about the overall textual arrangement, limited not so much by the expected attention span of the visitor but rather working against constraints of the exhibition context: the length of the experience and the number of available delivery devices determine the maximum number of spectators that can participate in the experience over time. Finding the right trade-off between giving enough space to explore and pacing the experience through an event-based system supported by stewards did the trick for us.

A mix of spatial triggers (such as being in the right location and looking at the right launch target) and time-based triggers was paramount. For example, in the first part of the experience, the posters and overlay videos serve as decoration, but the visitor is ushered through this experiential tunnel by the poem. As the poem concludes, it invites the user to move on to the next zone without any explicit announcement—each user was able to intuit the next stage of the experience using purely contextual cues.

The interactive element is important, too, personalizing the experience through path visualization (in version 1) and the compound creator (in version 2), thus making every experience unique and not just a standardized consumption. This seems particularly important in the context of using personal, body-worn devices and may be different for installations where the user is not wearing equipment, like with spatial projection where multiple persons observe an experience at the same time.

The aspect of co-creating and being in charge of the experience seems to be the big advantage of the chosen approach. This comes out rather weakly in the part of the experience where gaze triggers overlay videos on the posters but is particularly strong in effect where the user co-creates new words using the word composer.

This latter aspect is something that we intend to further investigate in the future, supported by a grant of the Independent Social Research Foundation (ISRF). We plan to investigate whether it is possible to push the envelope regarding co-creation, creating multiple layers of overlay poems rehashing words, syllables, and perhaps even letters from one part of the experience into another. Moreover, we are fascinated by the relationship between word and space, and we are keen to investigate whether it is possible to mimic the geometry of the space in both the arrangement and selection of words woven into a narrative provided by poetry.

This work does not come without limitations, which reach from the rather limited available resources for the product and the exhibition (such as the time constraints, development power, and financial support of the project), limited numbers of experiment participation, and limits of hardware and software. Limitations, however, were also found in the review of existing research. Little to no attention has been paid to word art and arts performance within the area of AR escape rooms, which this article seeks to help remedy.

Significantly, the impact of technology change in a fast-growing topical area results in fast turnover of hardware and software support, giving way to new technologies that are more effective and have greater performance to replace the old. As seen within the project, the development cycle of UNBODY was undertaken on a HoloLens 1, which, by Microsoft's standard as of 2020, has been retired to make way for the second generation, promising enhanced functionality, support, performance, and comfort.

From the research, the escapED framework was utilized and loosely followed in order to create the UNBODY design and implementation. Much of the framework provided the project with a focus on the main objectives to achieve throughout its cycle.

Despite the user evaluation of the first prototype of UNBODY being relatively successful in yielding results that could be applied to create a more refined prototype version of the project, limitations also lay within the methodology of the feedback data collection. The SPINE scale may need further refinement to analyze and conclude a broader picture for bigger applications.

Time constraints play a large limitation within any research project. Seen within UNBODY, the time frame to produce a product was small, so quick, agile sprints helped to promptly deliver working functionality. More time would have resulted in more functionality. Nevertheless, UNBODY was selected as a finalist for the 2020 Auggie awards in the category 'Best Art or Film.'

More research will undoubtedly be required in order to gain a deeper understanding of design frameworks and methodologies, principles, and considerations for building escape rooms dedicated to the spoken and written word. We hope that the open source project UNBODY could provide a small contribution to that, pushing and blurring the boundaries of technical and creative practice. **Supplementary Materials:** The following are available online, Video S1: https://vimeo.com/ channels/pal/394896384, UNBODY (Spatial Poetry).

Author Contributions: Conceptualization, F.W., L.M., J.B., E.W. and J.T.; methodology, L.M. and F.W.; software, L.M. and F.W.; validation, L.M. and F.W.; assets, J.B.; writing—original draft preparation, F.W. and L.M.; writing—review and editing, E.W.; visualization, J.B., F.W., L.M., J.T. and E.W.; supervision, F.W. and E.W.; funding acquisition, E.W. and F.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the European Commission under the Marie Skłodowska-Curie Programme as part of the CO-CREATION project (grant agreement no 734770) and under the Horizon 2020 Programme (H2020) as part of the ARETE project (grant agreement no 856533).

Institutional Review Board Statement: No review by ethics board required due to low-risk nature of the project (by check list E1).

Informed Consent Statement: No informed consent required, questionnaire data gathered anonymous with no personal characteristics.

Data Availability Statement: Data available from the authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Brooke, J. SUS: A Retrospective. J. Usability Stud. 2013, 2, 29–40.
- 2. Wild, F.; Vovk, A.; Guest, W. Evaluating Augmented Reality. 2020. Available online: https://www.overleaf.com/read/ pycrkjkjsbgt (accessed on 22 July 2021).
- Milgram, P.; Takemura, H.; Utsumi, A.; Kishino, F. Augmented reality: A Class of Displays on the Reality-Virtuality Continuum. In Proceedings of the Telemanipulator and Telepresence Technologies, Boston, MA, USA, 21 December 1995; Witmer & Singer: Orlando, FL, USA, 1994; Volume 2351, pp. 282–292.
- 4. Wild, F.; Perey, C.; Hensen, B.; Klamma, R. *IEEE Standard for Augmented Reality Learning Experience Models*; Institute of Electrical and Electronics Engineers (IEEE): Manhattan, NY, USA, 2020; pp. 1–3.
- 5. Azuma, R.T. A Survey of Augmented Reality. Presence Teleoper. Virtual Environ. 1997, 6, 355–385. [CrossRef]
- Speicher, M.; Hall, B.D.; Nebeling, M. What is Mixed Reality? In Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI'19), Glas-gow, Scotland, UK, 4–9 May 2019; pp. 1–15. [CrossRef]
- 7. Zhou, Z.; Cheok, A.D.; Qiu, Y.; Yang, X. The Role of 3-D Sound in Human Reaction and Performance in Augmented Reality Environments. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* 2007, 37, 262–272. [CrossRef]
- 8. Nicholson, S. Peeking Behind the Locked Door: A Survey of Escape Room Facilities. 2015. White Paper. Available online: http://scottnicholson.com/pubs/erfacwhite.pdf (accessed on 22 July 2021).
- 9. Nicholson, S. Creating Engaging Escape Rooms for the Classroom. Child. Educ. 2018, 94, 44–49. [CrossRef]
- Kim, T.-J.; Huh, J.-H.; Kim, J.-M. Bi-directional education contents using VR equipments and augmented reality. *Multimed. Tools Appl.* 2018, 77, 30089–30104. [CrossRef]
- 11. Clarke, S.; Peel, D.J.; Arnab, S.; Morini, L.; Keegan, H.; Wood, O. EscapED: A Framework for Creating Educational Escape Rooms and Interactive Games to For Higher/Further Education. *Int. J. Serious Games* **2017**, *4*, 73–86. [CrossRef]
- 12. Klamma, R.; Sous, D.; Hensen, B.; Koren, I. Educational Escape Games for Mixed Reality. In *Transactions on Petri Nets and Other Models of Concurrency XV*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2020; pp. 437–442.
- Warmelink, H.; Mayer, I.; Weber, J.; Heijligers, B.; Haggis, M.; Peters, E.; Max Louwerse, M. AMELIO: Evaluating the Teambuilding Potential of a Mixed Reality Escape Room Game. In *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play*; ACM: New York, NY, USA, 2017; pp. 111–123. [CrossRef]
- Pan, R.; Lo, H.; Neustaedter, C. Collaboration, Awareness, and Communication in Real-Life Escape Rooms. In Proceedings of the 2017 Conference on Interaction Design and Children, Palo Alto, CA, USA, 27–30 June 2017; ACM: New York, NY, USA, 2017; pp. 1353–1364.
- 15. He, L.; Li, H.; Xue, T.; Sun, D.; Zhu, S.; Ding, G. Am I in the Theater? In Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology, Tokyo, Japan, 28 November–1 December 2018; ACM: New York, NY, USA, 2018; p. 28.
- Schwelling, E.; Yoo, K. Automatic 3D Modeling of Artwork and Visualizing Audio in an Augmented Reality Environment. In Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology, Tokyo, Japan, 28 November–1 December 2018; ACM: New York, NY, USA, 2018; p. 47.
- 17. Plecher, D.; Ludl, M.; Klinker, G. Designing an AR-Escape-Room with Competitive and Cooperative Mode. In *GI VR/AR Workshop*; Weyers, B., Lürig, C., Zielasko, D., Eds.; Gesellschaft für Informatik e.V.: Bonn, Germany, 2020. [CrossRef]
- Zeng, H.; He, X.; Pan, H. Implementation of Escape Room System Based on Augmented Reality Involving Deep Convolutional Neural Network. In *Virtual Reality*; Springer: Berlin/Heidelberg, Germany, 2020. [CrossRef]

- 19. Vicari, C. Escape the Planet: Empowering Student Designers to Create a Science-Based Escape Room with Augmented Reality. *Int. J. Des. Learn.* **2020**, *11*, 80–95. [CrossRef]
- 20. Estudante, A.; Dietrich, N. Using Augmented Reality to Stimulate Students and Diffuse Escape Game Activities to Larger Audiences. J. Chem. Educ. 2020, 97, 1368–1374. [CrossRef]
- Figueroa, P.; Guo, R.; Takashima, K.; Weyers, B. Escape Room in Mixed Reality: 10th Annual 3DUI Contest. In Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, 23–27 March 2019; Institute of Electrical and Electronics Engineers (IEEE): Manhattan, NY, USA, 2019; pp. 1407–1408.
- 22. Davari, S.; Li, Y.; Lisle, L.; Lu, F.; Zhang, L.; Blustein, L.; Feng, X.; Gabaldon, B.; Kwiatkowski, M.; Bowman, D.A. Save the Space Elevator: An Escape Room Scenario Involving Passive Haptics in Mixed Reality. In Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, 23–27 March 2019; Institute of Electrical and Electronics Engineers (IEEE): Manhattan, NY, USA, 2019; pp. 1405–1406.
- Lopes, P.; You, S.; Ion, A.; Baudisch, P. Adding Force Feedback to Mixed Reality Experiences and Games Using Electrical Muscle Stimulation. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; ACM: New York, NY, USA, 2018; p. 446.
- 24. Gómez-Cambronero, Á.; González-Pérez, A.; Miralles, I.; Casteleyn, S. Mixed Reality Escape Room Video Game to Promote Learning in Indoor Environments. In *Edulearn19 Proceedings*; IATED: Burjassot, Spain, 2019; pp. 5857–5864.
- 25. Jing, K.; Nygaard, N.; Tanenbaum, T.J. Magia Transformo. In *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play;* ACM: New York, NY, USA, 2017; pp. 421–429.
- 26. Bernard, J. The Red and Yellow Nothing; Ink, Sweat and Tears Press: London, UK, 2016.
- 27. Bernard, J. Other Ubiquities; Ink, Sweat and Tears Press: London, UK, 2017.
- 28. Bernard, J. Surge, Chato; Ink, Sweat and Tears Press: London, UK, 2019.
- 29. Brooke, J. SUS: A Quick and Dirty Usability Scale. In *Usability Evaluation in Industry*; Jordan, T., Weerdmeester, M., Eds.; Taylor & Francis: Abingdon, UK, 1996; p. 189ff.
- Sauro, J. Measuring Usability with The System Usability Scale (SUS). 2011. Available online: https://measuringu.com/sus/ (accessed on 22 July 2021).
- 31. Sauro, J. A Practical Guide to Measuring Usability; Measuring Usability LLC: Denver, CO, USA, 2013; ISBN 978-1-4538-0656-2.
- Sauro, J. 10 Things to Know About the System Usability Scale (SUS). 18 June 2013. Available online: https://measuringu.com/10 -things-sus/ (accessed on 22 July 2021).
- BlinkUX. Usability Sample Size Calculator. Available online: https://assets.blinkux.com/usability-sample-size/ (accessed on 22 July 2021).
- Begnigno, F.; Ferreira, M.; Gomara, B.; Leonardi, L.; Matesanz, M.; Bertz, I.; Montagnino, F. Re-FREAM Art Tech Toolbox. 2020. Available online: http://www.re-fream.eu/wp-content/uploads/2020/07/Re-Fream_Art-Tech_Toolbox-v2-2020.pdf (accessed on 22 July 2021).
- 35. Krzywinska, T.; Phillips, T.; Parker, A.; Scott, M.J. From Immersion's Bleeding Edge to the Augmented Telegrapher. J. Comput. Cult. Herit. 2020, 13, 1–20. [CrossRef]