Journal of
Virtual Worlds Research
jvwr.net ISSN: 1941-8477

Pedagogy

Taking Stock and Looking Forward

January 2019 (Part 2)
Volume 12 No. 1

Cover: Photo by Josh Sorenson from Pexels

Volume 12, Number 1
Pedagogy - Taking Stock and Looking Forward
Part 2

January 2019

Cover image: Photo by Josh Sorenson from Pexels

Journal of Virtual Worlds Research
jvwr.net ISSN: 1941-8477

Volume 12, Number 1
Pedagogy – Taking Stock and Looking Forward (Part 2)
January 2019

Alternative Embodied Cognitions at Play: Evaluation of Audio-Based Navigation in Virtual Settings via Interactive Sounds

Enrico Gandolfi
Research Center for Education Technology, Kent State University
Robert Clements
Department of Biological Sciences, Kent State University

Abstract

The current study addresses the concept of mediated embodied cognition (EC) by focusing on audio inputs/outputs rather than sight and visual cues. N=10 subjects were involved with the navigation of five virtual mazes with different design and sensorial affordances. A virtual audiocane was developed (within a middleware - NEED - aimed to improve learning environments' accessibility) for supporting spatial exploration and scanning. Data collected span spatial thinking, previous knowledge of digital settings, completion time, ability to visualize virtual spaces, and usability and perception of the different design/sensorial situations provided. Results point to effective design choices and factors to consider when staging mediate EC instances. With these findings, the article addresses three current gaps in EC related literature: 1) lack of attention to special needs and *diverse* embodiments, 2) insufficient focus on audio inputs/outputs as interactive affordances; and 3) poor efforts in evaluating the design and assistive elements in EC integrations.

1. Introduction

Embodied Cognition (EC) can be defined as the involvement of multiple senses for enriching immersion and understanding (Clark, 2017; Mahon, 2015). Empirical evidence shows that digital approaches - and especially virtual reality - can ideally support this process pointing to increased learning outcomes (Shapiro, 2011). However, few studies have addressed this topic with an absence-based lens (i.e., removing one or more senses for challenging the others, and then creating an alternative embodied cognition) - which could be highly beneficial also because of its accessibility implications.

This study aims to fill this gap (and further related to EC literature) by testing an embodied-cognition related middleware called NEED (Necessity, Ego, Environment, Danger). This platform has been developed to facilitate embodied cognition and ideally support visually impaired as well as sighted students in exploring learning virtual environments. The program works as a layer giving additional sensorial cues (audio, haptic and beyond) triggered by features of the spatial environment to already existing interactive experiences for staging an inclusive multi-sensorial engagement and evaluating its utility for improving users' spatial and navigational awareness. Our investigation has addressed the audio component by involving n=10 subjects and related embodiment (Johnson-Glenberg, Birchfield, Tolentino, & Koziupa, 2014) in different conditions. The emphasis on blindness (in this case, artificial) is because this major disability creates fundamental boundaries directly affecting a person's' ability to perceive their surroundings and gather information. Moreover, it challenges sighted individuals to re-think their spatial and navigation skills. Participants have been engaged with spatial reasoning tasks including maze navigation and environmental interaction and understanding via different configurations of audio cues. Single performances and usability-related data were collected for understanding strengths and weaknesses of different solutions in terms of design and pedagogy - (e.g., use of one or multiple assistive modules within the same interface, the presence of introduction and assistance). Results suggest a list of guidelines and highlights to consider when adopting mediated embodied cognition sessions, with an emphasis on diversity and alternative viewpoints. Although the NEED platform has functioned as a concrete *reference ground* for the study, this article's implications can be extended to other tools and technologies with similar objectives and scopes. Aside from being pertinent to the topic of embodied cognition, the study's emphasis on navigation and spatial reasoning is motivated by the fact that understanding spatial properties such as shape, size, distance, orientation, and relative location is essential in the development of skills, environmental awareness and knowledge acquisition in many topics (from science to engineering). The article's desired outcome is to shed light on malleable factors in integrating embodied cognition into instruction and special education, which is a still overlooked topic that requires more analytical efforts and reflections.

2. Embodied Cognition(s)

Embodied cognition refers to the idea that our cognitive schemes are embodied rather than abstract dynamics. Body and senses are no more peripheral components of our thinking but rather structuring dimensions to consider spanning material environment, situation, and timing (Barsalou, 2010; Clark, 2017; Shapiro, 2011). Although its origins can be traced back to Gibson (1979) and Vygotsky (1978), the term has acquired increasing popularity in the last decades also because of its relation with technology spanning virtual reality and smart interfaces (Amin, Jeppsson, & Haglundm, 2015; Mahon, 2015). There are several approaches to embodied cognition, from support to criticism. The six core traits of EC suggested by Wilson (2002) are among the leading references for framing the topic. To summarize, EC implies that cognition is situated in real life (1); is influenced by time and related processes (2); engages with symbolic, social and, material affordances with the context of action (3); is extended and beyond the distinction between internal and external attributions (4); is driven by the goal to take action in the real world rather than representing it (5), and is shaped by body-based variables and constructs even when it is inner (6). As observed by Amin et al. (2015), for Wilson the last claim is the leading one because it entails that embodiment and sensorial experiences may have significant learning potential (for instance, triggering conceptual metaphors).

As argued by Bailey, Bailenson, and Casasanto (2016), there is empirical evidence that our mind tries to replicate and mimic grounded experiences with concrete outcomes, from knowledge gains to emotions. Pursuing this line, current literature tells us that body involvement may benefit instruction (and especially STEM fields) by making abstract concepts concrete and tangible (Alibali & Nathan, 2012, Dackermann; Kontra, Lyons, Fischer, & Beilock, 2015; Dunleavy & Dede, 2014; Gallagher & Lindgren, 2014; Glenberg, Witt, & Metcalfe, 2013; Soler, Contero & Alcañiz, 2017). Moreover, addressing multiple senses facilitates the involvement with mediated environments and settings (Maister, Slater, Sanchez-Vives, & Tsakiris, 2015; Romano, Llobera, & Blanke, 2016; Won, Bailenson, Lee, & Lanier, 2015). New technologies are ideally supporting this trend by providing multi-modal experiences and types of immersion, from multimedia and haptic media to augmented reality and head-mounted displays (Ferdig, Gandolfi, & Zimmel, 2018; Soler, Contero & Alcañiz, 2017; Yee, Bailenson, & Ducheneaut, 2009). Regardless, there are also several critical approaches to EC, which point to its vague definitions and implications and argue that EG is not a turning point in educational psychology as several authors believe (Goldinger, Papesh, Barnhart, Hansen, & Hout, 2016; Hommel, 2015). It can be argued that EC refers to a broad range of discussions with multiple disciplines at stake, from evolutionary theories to anti-cognitivist approaches (Hommel, 2015; Wilson & Golonka, 2013) spanning mild and radical embodiment theories and arguments. Technologies make this scenario even more complex to explore due to the novelty of related innovations (Ferdig et al., 2018; Freina & Ott, 2015).

3. Framing Embodied Challenges

The aim of this article is not to intervene in such a wide debate but rather to address three of its current gaps; in doing so, the hope is to provide suggestions for harnessing and reflecting on EC aside from the specific viewpoint adopted.

The first issue regards special needs, which are still an overlooked field in EC literature. Although there are several studies about assistive technology at large, EC has been rarely addressed in these terms. However, individuals with body impairments have a specific relation with embodiment (and especially with technology) because of the cognitive effort required to deal with their condition (Anderson, 2003). Addressing an absence can entail an alternative perspective on and within mediated and non-mediated experiences (Gandolfi, 2017), satisfying and even challenging the need to provide multiple expressive channels in education (CAST, 2018; Meyer, 2002). In other words, Alternative Virtual Environments (AVE) can provide information through a sensorial limitation that promotes the use of specific senses and embodiments (exploring fourth and sixth EC traits according to Wilson, 2002).

The second one relies on the fact that haptic/tactical feedback is usually the core output directly associated with virtual actions in EC studies; however, audio can be harnessed as well, and in a more synergic way than its usual implementation. Moreover, harnessing this sense means that more accessible technologies can be adopted. Haptic devices (e.g., gloves, Oculus Touch) are usually out of reach of the majority of educators and individuals, while better audio input/output systems can be easily embedded in a variety of digital settings - from learning virtual environments (e.g., Second Life, AltSpaceVR) to digital games (e.g., Overwatch, Fortnite, League of Legends). In other words, there is a partial adoption of the so-called "multimedia principle," which states (Meyer, 2002; Sorden, 2012) that the presence of multiple senses can support learning by addressing different processing channels.

Such attention may deepen the fourth EC trait advanced by Wilson (2002), going beyond the difference between internal and external attributions.

The third issue is the scarcity of design and pedagogical reflections related to EC technological integration. Embodied interactions seem to work (or not) aside from a) how the sensorial feedback is planned and handled and b) users' previous knowledge. On the contrary, how inputs are given and the users' background can make a crucial difference in how virtual experiences are perceived and evaluated. In other words, setting the stage for a mediated embodiment does not work by itself, but it also depends on how sensorial inputs are handled and provided (third and sixth EG traits suggested by Wilson, 2002; see the so-called T-PACK model; Koehler & Mishra, 2009).

The aim of this article is to address the aforementioned three issues by removing one of leading elements of EC - i.e., vision (Barsalou, 2010; Shapiro, 2011; Weisberg & Newcombe, 2017) - and testing the design and implementation of an audio input/output system within the virtual setting. Our objective is to understand what are the best affordances in designing an effective audio embodiment by removing the sight from the equation. In this way, we also intend to enlighten how digital environments can be made more inclusive by strengthening the interplay between audio component and interaction.

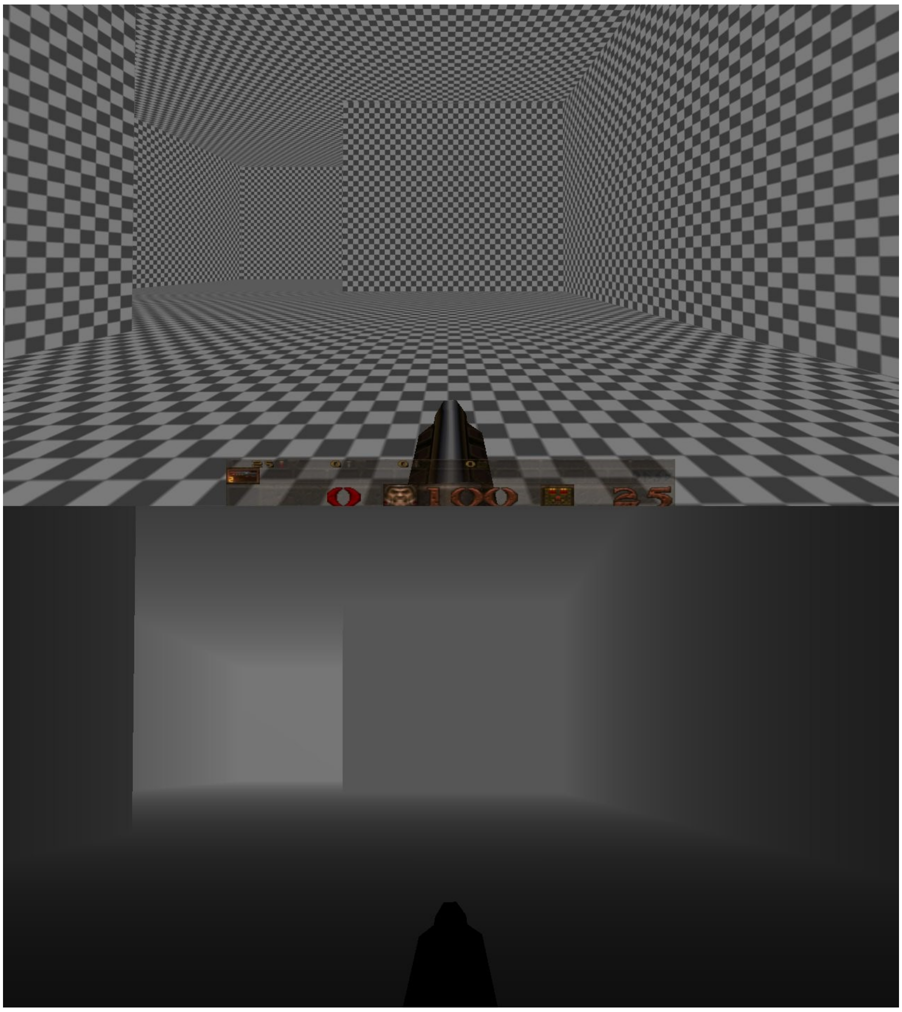
The effectiveness of the present study relies on the concept of spatial thinking, which is ideally related to EC - being and inhabiting a space as a fundamental cognitive process. Spatial thinking refers to the ability to solve space-related problems and exploit spatial information for reaching specific goals. It has been widely studied as a crucial skill to develop and consider while teaching, involving areas like architecture, mathematics, navigation, and critical thinking itself (Bednarz & Lee, 2011; McHarg, 1995; Tomaszewski, Szarzynski & Schwartz, 2014). In the present study, we refer to the overview provided by the National Research Council (2006; see also Lee & Bednarz, 2012), which breaks ST in three different criteria - how space is framed, how space is represented, and how space is debated.

4. Research Design

According to these premises, a middleware (NEED) was developed to improve audio (and potentially haptic) inputs within virtual navigation. The consequent study tested NEED with subjects whose eyes covered. This sensorial limitation was planned to address the first aforementioned gap (creating then an AVE instance) and not to interfere with the audio evaluation (second gap), which was actually improved. The evaluation addressed ST by adopting virtual mazes that the subjects were asked to navigate and finish according to different design solutions and supports (third gap). This specific type of challenge was chosen because it requires spatial thinking and reasoning (the objective is to understand and navigate an environment and its spatial relations) and grounds virtual embodied cognition via a specific timed task (second and fifth traits suggested by Wilson, 2002). As mentioned above, the main NEED feature used for the present study is an audio cane that is aligned with user aim while exploring virtual settings; its function is to provide a *sensorial sonar* that can be used for framing and monitoring space even without sight (e.g., in an AVE). Therefore, users have the opportunity to deal with space in novel ways. This interactive audio component has been inspired by blind canes for visually impaired individuals, which give non-visual outputs for scanning space as well.

4.1. The NEED Middleware - a Technical Overview

To improve the audio support, we developed a middleware system using open source and freely available tools. The software implementation functions by providing access to the depth buffer (see *Figure 1*) from the rendered game component and acquires pixel intensity data from the central region of the screen every 50 ms. The intensity of the grayscale pixel is subsequently used to control the pitch of an emitted beep. Using this method (from now on, audiocane), as the player navigates the maze, the distance of objects in the central portion of the screen controls the pitch of the emitted beep where low-frequency sounds indicate close objects while high-frequency beeps indicate objects far away. This solution was designed to associate the correct focus (absence of obstacles) with the most overwhelming feedback (high frequency), keeping the user on task. Our method utilized three components to perform these tasks and provide the audio output to the user. Interaction with the maze was provided by Quake Epsilon, a graphically enhanced version of the shareware game Quake (see box[1]) NEED's aim is to give an instrument to read mediated environments by using sensorial inputs beyond sight. The audiocane (the main instrument used in the present study) is just one of the NEED tools, which can also harness haptic feedback as well.



[1] Access to the depth buffer was acquired using ReShade a post-processing injector for games and video developed by crosire that provides access to additional video processing in binary applications as well as exposing frame color and depth information. Screen pixel intensity was probed and an associated beep generated by a custom written script using AutoHotkey an automation scripting language for Windows.

Figure 1: Depth buffer – see below

Figure 1. Depth buffer information was probed to identify distance of objects in the center of the screen from the user. The top image is the direct rendering of the maze and the lower image depicts the depth buffer from the same scene. Grayscale value is proportional to the distance from the user and is used to modify the pitch of the emitted beep.

4.2. Recruitment and Methods

The research took place in September and October 2018 in authors' university settings. Subjects were recruited via institutional university mailing lists under the supervision of authors' university I.R.B .committee. N:10 participants agreed to be involved with the study. The research design relied on the blinded (with a blindfold and/or keeping eyes closed according to subjects' preference) exploration of five virtual mazes[1] (see *Figure 2*). The navigation was handled by participants with four buttons (moving forward, moving backward, turn left, turn right). All the mazes included a checkpoint in the middle of the right path to reach the end. Checkpoints were characterized by an environmental sound (i.e., the sound increases with the closeness to the targeted location) for communicating the player his/her proximity to it. However, they differ according to the following features:

The first part (introduction and haptic manipulation)

1. 1A maze: easy maze with checkpoint audio
2. 1B maze: easy maze with checkpoint audio and preliminary exploration of a physical model of the maze itself (made with Lego blocks).

The second part (audiocane and environmental audio)

3. 2A maze: medium maze with checkpoint audio
4. 2B maze: medium maze with NEED audiocane and checkpoint audio (low volume)
5. 2C maze: medium maze with NEED audiocane and checkpoint audio (high volume)

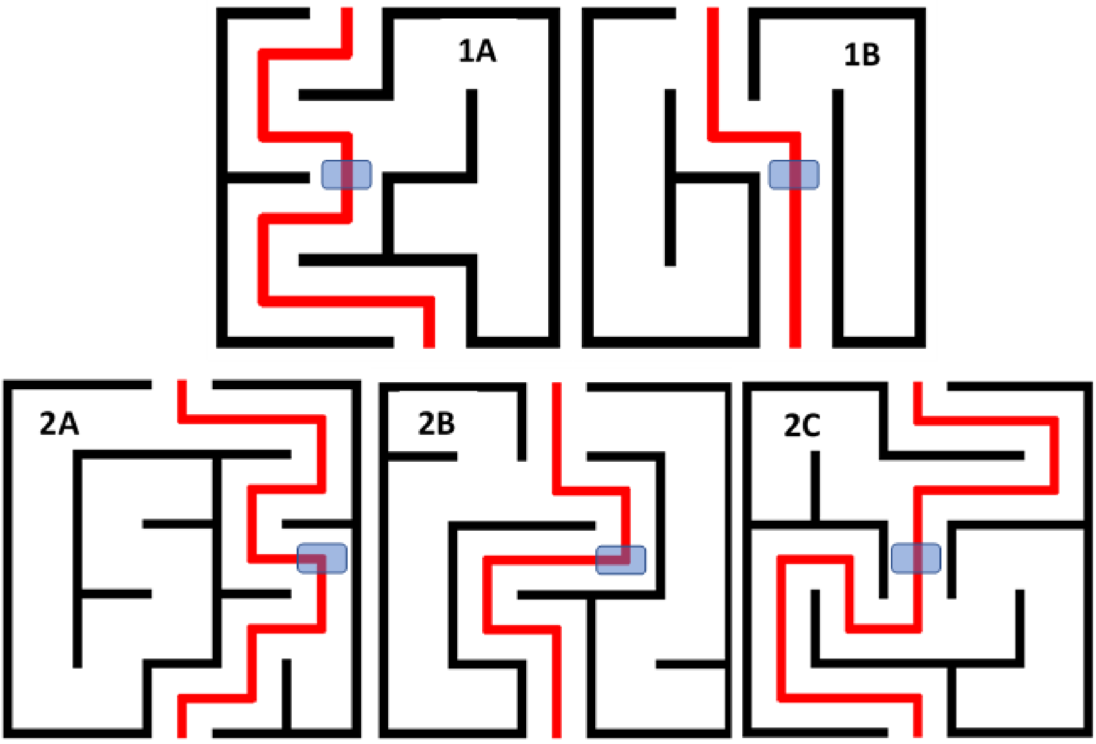


Figure 2: Top-down view

Figure 2. Images above depict a top-down view of each generated maze with solutions outlined in red used (from bottom to top). The top row is 4X4 orthogonal mazes while the bottom three are 5X5 orthogonal mazes. Transparent rounded blue rectangles indicate audio checkpoints for each maze and the starting point was located in the top entrance to the maze.

[1] Maze maps were generated with the online tool Maze Generator (http://www.mazegenerator.net/)

This progression was planned to 1) begin with a default situation (game sounds) (1A, 2A) - 2) add a haptic introduction to test potential influences (1B) - 3) reflect on the inclusion of the audiocane as a way to improve spatial recognition in dealing with standard (2C) or lower game sounds (2B) - this combination can affect cognitive load and therefore the ability to understand the environment itself (Sweller, 1988; Tran, Smith, & Buschkuehl, 2017)

The assessment relied on:

- Initial assessment in terms of gaming experience (i.e., previous experiences with virtual environments similar to Quake - yes/no) and spatial visualization ability (the ability to mentally visualize and manipulate 2D and 3D objects) using the Paper Folding Test – VZ-2[2] (Mitchell & Kent, 2003). The goal was to target previous knowledge in terms of technology and spatial thinking predisposition (from now on, ST) that might affect the experience.
- Time (find the exit) for each maze

Ability to reconstruct the path done (drawing or building it) for each maze just after its completion

- Evaluation: a five-item 1-6 Likert scale (from 1=I strongly disagree to 6=I strongly agree) about accessibility, informational outcome in comparison with alike (non-AVE) digital settings, easy-to-learn, potential translation into other virtual environments, perceived utility) about the audio cane with a text entry for additional comments.

From the performance itself, two scores were generated:

- Reconstruction Score: the skill to mimic the environment experienced. It is subdivided in:
    - Orientation - the ability to represent the overall direction of the path.
    - Self-Tracking - the ability to reproduce the completed path.
    - Environment - the ability to report walls and obstacles encountered during the experience.

Each component was made measurable from low (1) (0%-33% similarity between virtual experience and drawing) to high (67%-100% similarity between virtual experience and drawing), with (2) (34%-66% similarity between virtual experience and drawing) as medium score; the overall score was measured as a sum of three sub-dimensions, which can be considered different instances of space visualization. This skill concerns how the space is framed and represented.

- Exploratory behavior: the general attitude adopted during the navigation of virtual maze. It can be defined as methodical - i.e., the application of constant spatial patterns for framing the virtual environment - or random - i.e., the absence of strategies in reading the environment. This ability refers to how the space is debated and challenged.

Both the outcomes were measured through participant observation of the performance itself with an emphasis on spatial relations and patterns of navigation (Mark & Egenhofer, 1994). In addition, subjects' thinking aloud (Hoonhout, 2008) instances were gathered during the testing, and performance video recording and related heat maps were used as additional instruments for facilitating path interpretations via triangulation (Denzin, 2006).

This set of metrics aimed to address the following research question:

Research Question 1 (RQ1): do previous technological experience and spatial thinking affect performance in AVEs?

Moreover, results were used to test the hypothesis advanced below:

[2] The test is composed by 20 problems involving spatial rotation/visualization to solve in six minutes.

Hypothesis 1 (H1): the audiocane should improve the reconstruction score and facilitate a methodical approach due to its instrumental and inquiry function.

In addition, half of the subjects (#2,5,7,8,9) were positively but mildly coached (supportive commentary without additional information) (Renton, 2009) by researchers during the maze exploration - the coaching was limited to a known set of interactions, among which: (arriving at the checkpoint (e.g., "you did it"), using the audiocane in a methodical way (e.g., "nice job"), and encouragement when subjects were stuck (e.g., "it is alright, keep trying"). This well-established group of parameters was planned to give support without giving excessive advantages to the coached subjects. This difference was due to understanding external influence in a context with potentially overwhelming sensory inputs. The aim was to better comprehend if alternative embodied experiences may be affected by additional audio feedback from outside the screen (envisioning possible supervisions by instructors and peer mentors).

Research Question 2 (RQ2): What is the influence of supportive assistance in exploring AVEs? In other words, does it facilitate the experience or fuel an input overflow?

Finally, the second step (1B) was included for probing if manipulating the physical model of a virtual environment could facilitate its exploration (**Research Question 3 - RQ3**). T-Tests (two-tails) were directed for comparing reconstruction scores across different approaches.

5. Results

Table 1 reports age, gender, previous game expertise (GE), spatial thinking (ST) and timing spent for completing each maze (the exploration was stopped after 7 minutes). The participants recruited were n=6 females and n=4 males of different ages (m:31.2; st: 15.246). Half sample had previous experiences with gaming and virtual environments, while the other was novel to the technology adopted. Subjects #3,5,7,8,9 were not coached, while the others were. Three participants (#3,4,6) showed a high ST score in comparison with the rest of the sample.

Table 1: Results - Part One

Results - part one							
#	**GE**	**ST**	**1A**	**1B**	**2A**	**2B**	**2C**
1	yes	4	3:43	2:05	4:03	1:42	6:12
2	no	4	0:32	4:25	1:40	4:00	+7m
3	yes	19	3:01	2:11	1:16	5:11	+7m
4	no	18	3:33	0:12	+7m	+7m	6:47
5	no	8	2:31	1:00	3:39	4:06	+7m
6	yes	15	4:45	0:16	1:10	6:50	4:53
7	yes	5	2:25	4:43	2:39	4:31	5:02
8	no	6	2:27	+7m	2:52	5:28	6:40
9	yes	3	1:00	0:45	2:01	6:16	2:29
10	no	7	5:35	2:30	2:12	2:59	+7m

Table 2 gives an overview of exploratory behavior (ExBeh) and reconstruction score (ReSco) by subject, while Table 3 spans percentages (of the leading ExBeh) and aggregate means (ReSco) by specific mazes and subgroups (GE versus non-GE, high ST versus low ST, absence/presence of coaching).

Table 2: Results - Part Two

Results - part two		
#	**ExBeh**	**ReSco (1A/1B/2A/2B/2C)**
1	Methodical 1B/2B/2C Random 1A/2A	6/9/5/8/9
2	Methodical 2C Random 1A/1B/2A/2B	5/4/3/5/8
3	Methodical 1B/2B/2C Random 1A/2A	6/8/4/7/9
4	Methodical 1B/2B Random 1A/2A/2C	5/9/3/9/9
5	Methodical 1B/2C Random 1A/2A/2B	3/6/4/6/7
6	Methodical 1B/2A/2B/2C Random 1A	5/8/6/9/9
7	Methodical 1B/2B/2C Random 1A/2A	4/8/5/9/9
8	Methodical 2C Random 1A/1B/2A/2B	3/5/4/6/8
9	Methodical 1B/2A/2B/2C Random 1A	5/9/5/8/8
10	Methodical 1B/2B/2C Random 1A/2A	4/6/5/7/8

Table 3: Results - Part Three

Results - part three						
Measure		**1A**	**1B**	**2A**	**2B**	**2C**
ExBeh(%)		Random 100%	Methodical 80%	Random 80%	Methodical 70%	Methodical 90%
	GE(%) nonGE	Random 100% Random 100%	Methodical 100% Methodical 60%	Random 60% Random 100%	Methodical 100% Random 60%	Methodical 100% Methodical 80%
	hST(%) lST(%)	Random 100% Random 100%	Methodical 100% Methodical 71.42%	Random 66.66% Random 85.71%	Methodical 100% Methodical 57.14%	Methodical 100% Methodical 85.71%
	CO(%) nonCO(%)	Random 100% Random 100%	Methodical 80% Methodical 80%	Random 80% Random 80%	Methodical 80% Methodical 60%	Methodical 80% Methodical 100%
ReSco(m)		4.6	7.2	4.4	7.4	8.4
	GE/nonGE(m)	5.2/4	8.4/6	5/3.8	8.2/6.6	8.8/8
	hST/lST(m)	5.33/4.28	8.33/6.71	4.33/4.42	8.33/7	9/8.14
	CO/nonCO(m)	5/4.2	7.2/7.2	4.4/4.4	7.6/7.2	8.6/8.2

Results point to increasing reconstruction scores and methodic attitudes through the experiment (**H1**). The least understood mazes were 1A/2A, whose audio system relied on game sounds. The haptic introduction to 1B seemed effective in setting the stage for spatial understanding (**RQ3**) (1A-1B: p-value< 0.005). The audiocane proved to be a valuable addition in navigating virtual mazes via audio. It supported maze visualization and a methodical attitude in almost all the subjects, who used it for mapping and tracking the digital environment they were asked to explore (2A-2B: p-value< 0.005; 2A-2C: p-value< 0.005). It can be argued that 2B served as an introduction to the audiocane, which was properly harnessed in 2C; however, no significant difference emerged between 2B and 2C (p-value> 0.005). In-game environmental audio volume did not represent a variable to consider in these two mazes (therefore, users can be let free to change it). Coaching appeared irrelevant in influencing users' experience, with no significant differences between the two groups (**RQ2**). Game expertise seems to support AVE navigations: gamers scored higher ReSco than non-gamers with a more prominent methodical approach. However and although this trend is glaring, differences were not significant; NEED seems to be accessible and work properly aside from the presence of GE. The impact of spatial thinking is more challenging to address due to the different size of the subgroups; however, it can be argued that the three subjects with high ST achieved better scores preferring a methodical approach (**RQ1**). Usability scores were good for accessibility (m:4.4; st:1.26) and information outcome (m:4.2; st:1.98) and high for easy-to-learn (m:5, st:1.41) , potential translation into other virtual environments (m:5.7; st:0.48), and perceived utility (m:5.6; st:0.96). Comments highlighted the novelty of the experience itself, the importance of the audiocane in exploring and understanding the environment (in comparison with mazes without it), the need to better differentiate checkpoints sounds from exit point ones, and the struggles with walls and corners in 1A,1B, and 2A. During the experience itself, all the subjects expressed their initial confusion (e.g., "wow, it seemed easy but...it is not easy at all…"), the unconventionality of the task (e.g., "this is the first time I am doing something like that"; "I am trying to build the walls in my head...this is new"), and the majority (n=8) the satisfaction in using the audiocane (e.g., "now I get it...I can understand; "yes, I am mapping this room").

6. Discussion

It can be argued that providing a sensorial instrument like the NEED audiocane benefits spatial exploration in AVEs, with interesting implications in terms of a) an alternative embodied cognition in virtual spaces and b) overall accessibility and customization. Although game expertise counts, the increasing ability to read the environment was widespread among the subjects, who present heterogeneous traits (different age, different ST and approach to virtual environments). This is even more significant considering that playing blindfolded was new to all participants and that no introduction was given about the audiocane (just a description of its function). The evaluations reiterate this point with promising scores in terms of learning pace and overall potential. It is interesting to notice the importance of having prior experiences with virtual environments and games (GE). It regards embodied memories that gaming subjects were able to re-enact even with not visual included. This sub-group was the first to use a methodical exploration starting from the simple – but not immediate - premise that there was a virtual environment to inhabit. Regardless, non-gamers increasingly followed this strategy with the audiocane, which worked as a triggering tool for staging spatial reasoning. This finding is aligned with other studies targeting the synergy between spatial thinking and video games at large (e.g., Choi & Feng, 2017; Uttal et al., 2013).

It was interesting that the presence of a coach did not make a concrete difference; the same highlight can be advanced about in-game sounds as an additional output for the audiocane. According to thinking aloud notes and comments, the audiocane itself become the leading tool to exploit; moreover, coach's prompts were considered *external* (therefore, marginal), outside the learning environment inhabited by the subjects.

7. Conclusions

The present study is, however, just the first step in better understanding alternative sensorial embodiments. Therefore, it is exploratory at its core with a small sample and a focus on emerging trends rather than statistical significance. Further investigations are required deploying different methods, interactive experiences, and audiences; for instance, visually impaired people will be involved in our follow up studies for uncovering their specific attitude in dealing with AVEs, which actually aim to be more inclusive virtual environments. Moreover, spatial thinking and gaming expertise are multi-faced concepts characterized by several proposals and frameworks, and additional instruments and theories can be aligned with the ones adopted in the study. Furthermore, the simple game elements and the exclusive focus on the audio component limit the scope of related findings; next studies should address more complex interactive environments and enlighten the haptic component better. Finally, the NEED audiocane itself follows a specific design that may be stressed, improved and personalized in future iterations (especially through haptic cues, which are already under development).

Aside from these limitations, the desired outcome of the current study is to have presented a novel approach to an alternative embodied cognition. By doing so, the objective was to address the three aforementioned gaps related to EC literature (special needs, audio, design, and implementation) and advance AVEs as an opportunity for re-thinking EC and its potential integration. Moreover, the implications can be noteworthy for practitioners and educators because the range of applications of our proposal is wide, from virtual learning environments to digital games. Although there is a strong emphasis on how innovative technologies (e.g., augmented reality, immersive virtual reality) can influence EC, desktop virtual reality (which is still the most popular one) keeps providing remarkable opportunities for creating meaningful embodied experiences and improving digital settings' accessibility and equality. Such an "update" can assist both active interaction (e.g., like in the present study) and passive media consumption (e.g., videos and streaming involving virtual environments, which can be made more intelligible via an interactive audio output like the NEED audiocane). This article's aim is to take the first step in realizing such a potential.

References

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences*, *21*(2), 247-286.

Amin, T. G., Jeppsson, F., & Haglund, J. (2015). Conceptual Metaphor and Embodied Cognition in Science Learning: Introduction to special issue. *International Journal of Science Education*, *37*(5-6), 745-758.

Anderson, M. L. (2003). Embodied cognition: a field guide. *Artificial Intelligence*, *149*(1), 91-130.

Bailey, J. O., Bailenson, J. N., & Casasanto, D. (2016). When Does Virtual Embodiment Change Our Minds?. *Presence*, *25*(3), 222-233.

Barsalou, L. W. (2010). Grounded cognition: Past, present, and future. *Topics of Cognitive Science*, *2*(4), 716-724.

Bednarz, R. S., & Lee, J. (2011). The components of spatial thinking: empirical evidence. *Procedia - Social and Behavioral Sciences*, *21*, 103-107.

CAST (2018). *Universal design for learning guidelines*. Retrieved from: http://www.udlcenter.org/aboutudl/udlguidelines.

Choi, H. S. & Feng, J. (2017). Using Video Games to Improve Spatial Skills. In R. Zheng & M. K. Gardner (Eds.), *Handbook of Research on Serious Games for Educational Applications* (pp. 93-104). Hershey, PA: IGI Global

Clark, A. (2017). Embodied, situated, and distributed cognition. In W. Bechtel & G. Graham (Eds.), *A Companion to Cognitive Science* (pp. 506-517). Hoboken, NJ: Wiley.

Dackermann, T., Huber, S., Bahnmueller, J., Nuerk, H. C., & Moeller, K. (2015). An integration of competing accounts on children's number line estimation. *Frontiers in Psychology, 15*(6), 1-7.

Denzin, N. (2006). *Sociological Methods: A Sourcebook*. Piscataway, NJ: Aldine Transaction.

Dunleavy, M., & Dede, C. (2014). Augmented reality teaching and learning. In M. D. M. J. M. Spector, J. Elen, & M. J. Bishop (Eds.), *The Handbook of Research for Educational Communications and Technology* (4th ed.) (pp. 735-745). New York, NY: Springer.

Ferdig, R. E., Gandolfi, E., & Immel, Z. (2018). Educational opportunities for virtual reality. In J. Voogt, G. Knezek, R. Christensen, & K-W Lai (Eds.), *Handbook of Information Technology in Primary and Secondary Education* (pp. 1-12). Cham, Switzerland: Springer International Publishing.

Foglia, L., & Wilson, R. A. (2013). Embodied cognition. *WIREs Cognitive Science, 4*(3), 319-325.

Freina, L., & Ott., M. (2015). A Literature Review on Immersive Virtual Reality in Education: State Of The Art and Perspectives. In I. Roceanu (Ed.), *Proceedings of eLearning and software for education (eLSE)* (pp. 133-141). Bucharest, Romania: Carol I NDU Publishing House.

Gallagher, S., & Lindgren, R. (2015). Enactive Metaphors: Learning Through Full-Body Engagement. *Educational Psychology Review, 27*(3), 391-404.

Gandolfi, E. (2017). Beyond Diagonal Sciences: Applying Roger Caillois's Concepts of Symmetry and Dissymmetry to Journey. *Games and Culture*, *12*(4), 361-380.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Goldinger, S. D., Papesh, M. H., Barnhart, A. S., Hansen, W. A., Hout, M. C. (2016). The poverty of embodied cognition. *Psychonomic bulletin & review, 23*, 959-978.

Glenberg, A. M., Witt, J. K., & Metcalfe, J. (2013). From the revolution to embodiment: 25 years of Cognitive Psychology. *Perspectives on Psychological Science, 8*(5), 573-585.

Hommel, B. (2015). The theory of event coding (TEC) as embodied-cognition framework. *Frontiers of Psychology, 6*, 1-5.

Hoonhout, H. C. M. (2008). Let the game tester do the talking: Think aloud and interviewing to learn about the game experience. In K. Isbister & N. Schaffer (Eds.), *Game usability* (pp. 65-77). Burlington, MA: Morgan Kaufmann.

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psycholog*y, *106*(1), 86-104.

Koehler, M., & Mishra, P. (2009). What is Technological Pedagogical Content Knowledge (TPACK)?. *Contemporary Issues in Technology and Teacher Education, 9*(1), 60-70.

Kontra, C., Lyons, D., Fischer, S., & Beilock, S. L. (2015). Physical experience enhances science learning. *Psychological Science, 26*, 737-749.

Lee, J., & Bednarz, R. (2012). Components of spatial thinking: Evidence from a spatial thinking ability test. *Journal of Geography, 111*(1), 15-26.

Mahon, B. Z. (2015). What is embodied about cognition? *Language. Cognition and Neuroscience*, *30*(4), 420-429.

Maister, L., Slater, M., Sanchez-Vives, M. V., & Tsakiris, M. (2015). Changing bodies changes minds: Owning another body affects social cognition. *Trends in Cognitive Sciences, 19*(1), 6-12.

Mark, D. M., & Egenhofer, M. J. (1994). Modeling Spatial Relations Between Lines and Regions: Combining Formal Mathematical Models and Human Subjects Testing. *Cartography and Geographic Information Systems*, *21*(4), 195-212.

Mitchell, J., & Kent, L. (2003). Mental rotation: What is it?. *International Journal of Human-Computer Studies, 49*(1), 59-78.

McHarg, I. (1995). *Design with nature* (25th Anniversary Edition). Hoboken, NJ: Wiley.

National Research Council (2006). *Learning to Think Spatially Spatially*. Washington, DC: National Academies Press.

Renton, J. (2009). Coaching and Mentoring: What They are and How to Make the Most of Them. New York, NY: Bloomberg Press.

Romano, D., Llobera, J., & Blanke, O. (2016). Size and viewpoint of an embodied virtual body affect the processing of painful stimuli. *The Journal of Pain, 17*(3), 350-358.

Shapiro, L. (2011). *Embodied cognition.* New York, NY: Routledge.

Soler, J. S., Contero, M., & Alcañiz, M. (2017). VR Serious Game Design Based on Embodied Cognition Theory. In M. Alcañiz, S. Göbel, M. Ma, M. Fradinho, O. Jannicke, B. Hauge & T. Marsh (Eds.), *Joint International Conference on Serious Games JCSG 2017: Serious Games proceedings* (pp 12-21).New York, NY: Springer.

Sorden, S. (2012). The cognitive theory of multimedia learning. In B. J. Irby, G. Brown, R. Lara-Alecio & S. Jackson (Eds), *Handbook of Educational Theories* (pp. 155-168). Charlotte, NC: Information Age Publishing.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257-285.

Tomaszewski, B., Szarzynski, J., & Schwartz, D. I. (2014). Serious Games for Disaster Risk Reduction Spatial Thinking. In A. U. Frank (Eds.), *Proceedings of the 11th International Conference of the International Association For The Study Of Information Systems For Crisis Response And Management (ISCRAM)* (PP. 1-5). Vienna, Austria: ISCRAM.

Tran, C., Smith, B., & Buschkuehl, M. (2017). Support of mathematical thinking through embodied cognition: Nondigital and digital approaches. *Cognitive Research: Principles and Implications, 2*(16), 1-18.

Uttal, D. H., Meadow, N.G., Tipton, E., Hand, L. L., Alden, A.R., Warren, C., & Newcombe, N. S. (2013). The Malleability of Spatial Skills: A Meta-Analysis of Training Studies. *Psychological Bulletin, 139*(2), 352-402.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Weisberg, S.M., & Newcombe, N. S. (2017). Embodied cognition and STEM learning: overview of a topical collection in CR:PI. *Cognitive Research: Principles and Implications, 2*(38), 1-6.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625-636.

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers of Psychology*, 4(58), 1-13.

Won, A. S., Bailenson, J., Lee, J., & Lanier, J. (2015). Homuncular flexibility in virtual reality. *Journal of Computer-Mediated Communication, 20*(3), 241-259.

Yee, N., Bailenson, J. N., & Ducheneaut, N. (2009). The proteus effect: Implications of transformed digital self-representation on online and offline behavior. *Communication Research, 36*(2), 285-312.