The Impact of 3D Virtual Environments with Different Levels of Realism on Route Learning

A Focus on Age-Based Differences

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)
vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Ismini Eleni Lokka
aus
Griechenland

Promotionskommission
Prof. Dr. Sara I. Fabrikant (Vorsitz)
Prof. Dr. Arzu Çöltekin (Leitung der Dissertation)
Prof. Dr. Robert Weibel
Prof. Dr. Jan M. Wiener

Zürich, 2020

Summary

With technological advancements, it has become notably easier to create virtual environments (VEs) depicting the real world with high fidelity and realism. These VEs offer some attractive use cases for navigation studies looking into spatial cognition. However, such photorealistic VEs, while attractive, may complicate the route learning process as they may overwhelm users with the amount of information they contain. Understanding how much and what kind of photorealistic information is relevant to people at which point on their route and while they are learning a route can help define how to design virtual environments that better support spatial learning. Among the users who may be overwhelmed by too much information, older adults represent a special interest group for two key reasons: 1) The number of people over 65 years old is expected to increase to 1.5 billion by 2050 (World Health Organization, 2011); 2) cognitive abilities decline as people age (Park et al., 2002). The ability to independently navigate in the real world is an important aspect of human well-being. This fact has many socio-economic implications, yet age-related cognitive decline creates difficulties for older people in learning their routes in unfamiliar environments, limiting their independence. This thesis takes a user-centered approach to the design of visualizations for assisting all people, and specifically older adults, in learning routes while navigating in a VE. Specifically, the objectives of this thesis are threefold, addressing the basic dimensions of:

- **Visualization type as expressed by different levels of realism**: Evaluate *how much* and *what kind* of photorealistic information should be depicted and *where* it should be represented within a VE in a navigational context. It proposes visualization design guidelines for the design of VEs that assist users in effectively encoding visuospatial information.
- **Use context as expressed by route recall in short- and long-term**: Identify the implications that different information types (visual, spatial, and visuospatial) have over short- and long-term route recall with the use of 3D VE designs varying in levels of realism.
- **User characteristics as expressed by group differences related to aging, spatial abilities, and memory capacity**: Better understand how visuospatial information is encoded and decoded by people in different age groups, and of different spatial and memory abilities, particularly while learning a route in 3D VE designs varying in levels of realism.

In this project, the methodology used for investigating the topics outlined above was a set of controlled lab experiments nested within one. Within this experiment, participants' recall accuracy for various visual, spatial, and visuospatial elements on the route was evaluated using three visualization types that varied in their amount of photorealism. These included an Abstract, a Realistic, and a Mixed VE (see Figure 2), for a number of route recall tasks relevant to navigation. The Mixed VE is termed "mixed" because it includes elements from both the Abstract and the Realistic VEs, balancing the

amount of realism in a deliberate manner (elaborated in Section 3.5.2). This feature is developed within this thesis. The tested recall tasks were differentiated based on the type of information being assessed: visual, spatial, and visuospatial (elaborated in Section 3.6.1). These tasks were performed by the participants both immediately after experiencing a drive-through of a route in the three VEs and a week after that; thus, addressing short- and long-term memory, respectively. Participants were counterbalanced for their age, gender, and expertise while their spatial abilities and visuospatial memory capacity were controlled with standardized psychological tests.

The results of the experiments highlight the importance of all three investigated dimensions for successful route learning with VEs. More specifically, statistically significant differences in participants' recall accuracy were observed for:

1) the **visualization type,** highlighting the value of balancing the amount of photorealistic information presented in VEs while also demonstrating the positive and negative effects of abstraction and realism in VEs on route learning;
2) the **recall type**, highlighting nuances and peculiarities across the recall of visual, spatial, and visuospatial information in the short- and long-term; and,
3) the **user characteristics**, as expressed by age differences, but also by spatial abilities and visuospatial memory capacity, highlighting the importance of considering the user type, i.e., for whom the visualization is customized.

The original and unique results identified from this work advance the knowledge in GIScience, particularly in geovisualization, from the perspective of the "cognitive design" of visualizations in two distinct ways: (i) understanding the effects that visual realism has—as presented in VEs—on route learning, specifically for people of different age groups and with different spatial abilities and memory capacity, and (ii) proposing empirically validated visualization design guidelines for the use of photorealism in VEs for efficient recall of visuospatial information during route learning, not only for short-term but also for long-term recall in younger and older adults.

Acknowledgements

This thesis was carried out in the Department of Geography of the University of Zurich and would not have been possible without the intellectual and emotional support of the following people:

- **Prof. Dr. Arzu Çöltekin**: as my main supervisor, Arzu has always been there for me as a true mentor, with her endless guidance, trust, and motivation. She has always supported me both as a researcher, but also as an individual. She has provided me with numerous opportunities, all of which have helped me grow.
- **Prof. Dr. Sara Irina Fabrikant**: Sara has always been a great inspirational source for me. With her clear perspective, she has always guided me to the right path and with her questions and feedback has helped me reflect and understand the greater implications of my work.
- **Prof. Dr. Robert Weibel**: Rob has always provided critical and well-thought feedback helping me see the larger picture of my thesis. Without his connection to the URPP "Dynamics of Healthy Aging", the recruitment of older people would be impossible.
- **Prof. Dr. Jan M. Wiener**: Jan stepped in this project at a crucial time for me and with his expertise, he has filled in all the missing pieces. His always thoughtful feedback has substantially increased the quality of the thesis, while his always encouraging attitude has given me great strength to move on.
- **Dr. Christina Röcke**: Christina provided her expert guidance in the design and execution of the experiment. Without her support, the data collection from the older age group would have been impossible.
- **Prof. Dr. Susanne Bleisch**: who provided great contacts to advertise my experiment and collect the required data from older participants.
- **Ms. Brigitte Sonderegger**: due to her patience and support, the recruitment process of the older age group went smoothly and with great success.
- **Mr. Stefan Räber**: the editor of Carto.news, who helped me recruit older participants by advertising the experiment.
- **Prof. Dr. Reinhard Gottwald**: from the "Society of the history of Geodesy", who helped me recruit participants by advertising the experiment.
- **All my GIUZ colleagues and friends**: especially, Annina, Sascha, André, Gianluca, Sara, Alžběta, Katharina, Kenan, Kai-Florian, and Halldor who have supported me through my GIUZ times and with whom I have spent so many great moments in and out of the Irchel campus.
- **All my friends** and especially Alexandra, Konstantina, Veta, Maria, Fotini, Ioanna, Sophia, Natasha, Katerina, Angelos, Marie, Sophie, and Michi, who have always been there for me.
- **All the young and older experimental participants**: who voluntarily take part in both parts of my study and made this project reality.
- **Miltos**: without him, I would not have made it.
- **My family**: my father Phil, my mother Jasmin, my sister Kelly, and our little Ophelie, I owe you everything.

Thank you all!

Table of Contents

List of Figures

List of Tables

Abbreviations

ANOVA:	Analysis of Variance
GIVA:	Geographic Information Visualization and Analysis
GIScience:	Geographic Information Science
High-MRT:	Participants with high scores in the mental rotation test
High-VSM:	Participants with high visuospatial memory scores
LOD:	Level of detail
LOR:	Levels of realism
Low-MRT:	Participants with low scores in the mental rotation test
Low-VSM:	Participants with low visuospatial memory scores
MMSE:	Mini-Mental State Examination
MRT:	Mental rotation test
RQ:	Research question
UCD:	User centered design
URPP:	University Research Priority Program
VE:	Virtual environment
VR:	Virtual reality
VSM:	Visuospatial memory
WHO:	World Health Organization

I. SYNOPSIS

1. Introduction

1.1 Motivation

"Jenny turned 65 last year and retired. She now has more time to take up the painting classes that she always wanted. Her first lesson is this afternoon at 5pm. It will be her first time going to the painting studio, which is in an unfamiliar area of the city. Because she dislikes interacting with the navigation assistant and wants to avoid distraction while driving, Jenny typically looks up the information ahead of time and memorizes the route, focusing on specific shops and buildings along the route that will help her recall it. Once she feels confident that she has memorized the route, she begins her journey to the painting studio. On her way there, she arrives at an intersection where she fails to recall which way to continue. She starts to feel uneasy, but then decides to take a left turn, in hope that she will soon identify the next point she could recall. As she drives, she realizes that nothing reminds her of the rest of the route she had memorized. She is now lost. Trying to stay calm, she decides to return to the intersection where the uncertainty began. The time has now passed, and she will probably be late for her first class. When at the intersection, Jenny checks the other two options, and decides to continue straight from there. Luckily, she's back on track! She now recognizes the next landmark and finally arrives just few minutes after 5pm. She is still quite distressed from her unpleasant navigational experience and has a feeling of uncertainty for the next time she will have to take the same route."

Jenny's story represents an everyday navigational task, citing some of the complexities and difficulties of navigation. What this story also highlights is the *failure of her navigational assistance* to provide her with the adequate information to efficiently learn the required route and avoid such unpleasant experiences. This experience is especially prevalent in the daily lives of older adults, who might be slowly losing their ability to independently navigate.

There is ample evidence to suggest that spatial confusion and disorientation in unfamiliar environments can signal unhealthy aging (Hort et al., 2007; Iachini, Iavarone, Senese, Ruotolo, & Ruggiero, 2009) and spatial memory declines before verbal and visual memory (Barbeau et al., 2004). With weakening spatial memory, older adults may struggle with cognitive load at a lower threshold than younger adults when executing spatial tasks (Kessels, Meulenbroek, Fernandez, & Olde Rikkert, 2010). Such issues put older adults' ability to learn new routes, and thus their independent everyday mobility and functioning, at great risk. Navigating and route learning especially in unfamiliar destinations can be a difficult task at any age (Montello & Sas, 2006), but getting older adds a few more levels of complexity to this crucially important task (Moffat, Zonderman, & Resnick, 2001). The significance of being able to navigate independently without the support of other individuals is very important for older people, as the feeling of insecurity when being lost in an unfamiliar environment is unpleasant for the individual and for the people around them (McShane et al., 1998), in addition to its being expensive for society (Turcotte, 2013). Therefore, being able to provide older people with the right tools for their route learning can be of great value to these individuals and to society. To mention a few examples, scholars in psychology study healthy and unhealthy aging in connection with cognition (EUCAS[1]), focusing on understanding the cognitive functions that are affected by aging, while also working on "brain training" concepts to improve cognitive aging (Cassarino & Setti, 2015; Millington, 2012; Walton, Mowszowski, Lewis, & Naismith, 2014). In architecture and urban planning, scholars incorporate concepts such as "age-friendly cities" (Plouffe & Kalache, 2010) and "urban aging" (Eurocities[2]) into their research agendas with the goal of building sustainable environments that consider older adults as well (Fozard, Rietsema, Bouma, & Graafmans, 2000). In the sparse existing literature in geography that focuses on aging, researchers examine mobility and activity as indicators of healthy aging (Fillekes, Giannouli, Kim, Zijlstra, & Weibel, 2019; Fillekes, Kim, et al., 2019; Isaacson, Wahl, Shoval, Oswald, & Auslander, 2016) or of broader demographic and economic implications (Goodman, Brewster, & Gray, 2004; Hodge, 2008).

Virtual environments (VEs) are commonly used in navigation studies because they simulate real-world navigation and enable researchers to create controlled conditions for experimentation at the same time (Foreman, 2010; Kinateder et al., 2014). By definition, virtual reality (VR) displays aspire to mimic reality through highly photorealistic representations (Bowman & McMahan, 2007). However, visual realism needs to be carefully considered from a design perspective. There are many grounds on which visual realism is appreciated and desired, but it is also a target of criticism. For example, representing objects and phenomena at the highest degree of resemblance to their real-world counterparts can make it easier for people to relate to it (Finlayson, Zhang, & Golomb, 2017). Furthermore, people like realistic displays more when compared to abstract variations (Smallman & John, 2005b; Smallman & Cook, 2011). However, individual preferences and performance in spatial tasks with visually realistic displays can be misaligned, which is especially pronounced for people with lower spatial abilities (Smallman & John, 2005a, 2005b). Due to this conflict, visual realism is an important topic to study in the context of navigation. Too much realism might lead to information overload, especially for older people who, on average, exhibit a decline in their spatial abilities and visuospatial memory capacity. In geovisualization, an increasing effort has been evident in recent decades to integrate user studies that examine visualization designs from a human-centric perspective (*e.g.,* Çöltekin, Heil, Garlandini, & Fabrikant, 2009; Griffin & Fabrikant, 2012; Olson, 1997; Roth et al., 2017). In such studies, participant characteristics are especially important (Griffin & Fabrikant, 2012), but it appears that the focus is often on evaluating a visualization's *design* rather than on the characteristics of the *user*. While there is strong interest in usability, in addition to an increasing awareness of the importance of customizing and personalizing visualization solutions (Jameson, 2008; Steichen, Ashman, & Wade, 2012; Steichen & Fu, 2019), what it means to design explicitly for a target user group is currently under-investigated. Specifically, the older age groups are often neglected, most likely due to the convenience of using university students as sample participants ("convenience sampling") in user experiments. When examining the relevant literature (elaborated in Chapter 2), it becomes evident that there is minimal understanding of how aging may affect a person's performance with geographic visualizations. Psychology research clearly demonstrates a decline in many cognitive abilities during healthy aging

[1] https://eucas.org/

[2] https://eurocities.eu/eurocities/working_groups/Urban-Ageing&tpl=home

(Park et al., 2002), but we do not have a precise understanding of the implications of aging in spatial knowledge acquisition using various geographic visualization displays.

It has been clear to cartography scholars and the related communities that to design visualizations that work well, one needs to examine the topic "holistically" (Brewer, 2015). Such an examination may be based on at least three fundamental dimensions, *design*, *context and users' goals,* and *audience* of the visualization (Brewer, 2015; Çöltekin, 2019). These three dimensions may answer the questions of: 1) *what* is being built and *how*; 2) *why, where,* and *when* a phenomenon takes place, i.e., under which conditions, and with what intentions the visualization is created; and 3) *for whom* the visualization is being built, i.e., what are the individual and group differences (Fabrikant & Lobben, 2009; Lloyd & Bunch, 2005). Examining these three dimensions together, one can investigate the research questions holistically to identify which factors are important in designing the geographic visualization display.

Within this project, the *context* is navigation, specifically spatial knowledge acquisition in passive route learning, the examined *design* factor is visual realism in 3D VEs, and the *audience* is older adults with younger adults as the control group. From a fundamental science perspective, this Ph.D. project explores how 3D visualization design, especially different levels of visual realism in VEs, impacts spatial knowledge acquisition in route learning tasks for people of different ages and cognitive abilities. From an applied science perspective, the project examines how we can help improve older adults' route learning to eventually enable them to better navigate, and thus improve their overall well-being. To achieve these goals, this thesis employs a set of controlled laboratory experiments nested in one, which considers the three dimensions as explained above (Figure 1).

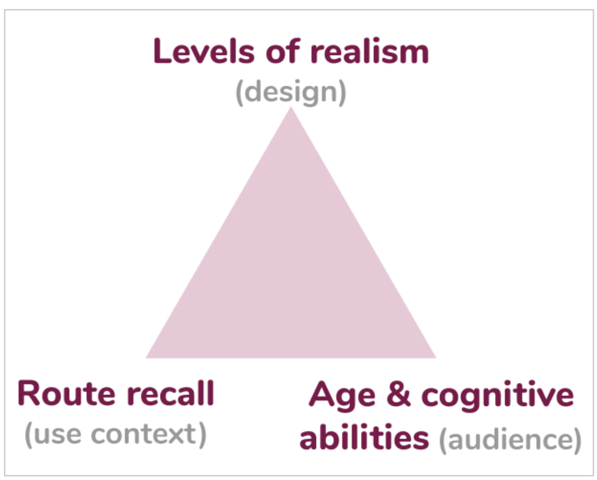


Figure 1*: Conceptual design of the project considering design, use context, and audience, framed in this thesis as levels of realism, route recall, and age & cognitive abilities respectively (recreated from Lokka & Çöltekin, 2016).*

Given the context above, the overarching question that guides this thesis is:

> *How can we design geographic visualization displays that enhance spatial knowledge acquisition for healthy aging older people in virtual route learning tasks to eventually assist them better in the complex task of navigation?*

1.2 Research questions

This dissertation is structured around the following three **leading research questions**, which tackle the three experimental dimensions described in the previous section (Section 1.1): (i) *visualization type/design*, expressed in this thesis as *levels of realism*, (ii) *information/task type* (use context), expressed in this thesis as *route recall in short- and long-term,* and (iii) *audience* (user characteristics) expressed in this thesis as *age differences and cognitive abilities*.

1.2.1 Visualization type: levels of realism

Terminology clarification: In the questions below, three *visualization types* are featured that correspond to the three VEs. Screenshots from these three VEs are shown in Figure 2, and their naming is briefly explained below (further elaborated in Section 3.5.2):

- ❖ **Abstract VE** serves as a baseline condition and includes a minimum amount of photographic information (Figure 2, left).
- ❖ **Realistic VE** is another baseline that simulates the real world, and the state-of-the-art VR representations of cities with a high amount of photorealism (Figure 2, right).
- ❖ **Mixed VE** results from a carefully designed combination of the Realistic and Abstract VEs, presenting photo-textures only on structures at critical locations for navigation ("effective landmark locations"), while the rest of the VE resembles the Abstract VE (Figure 2, middle).

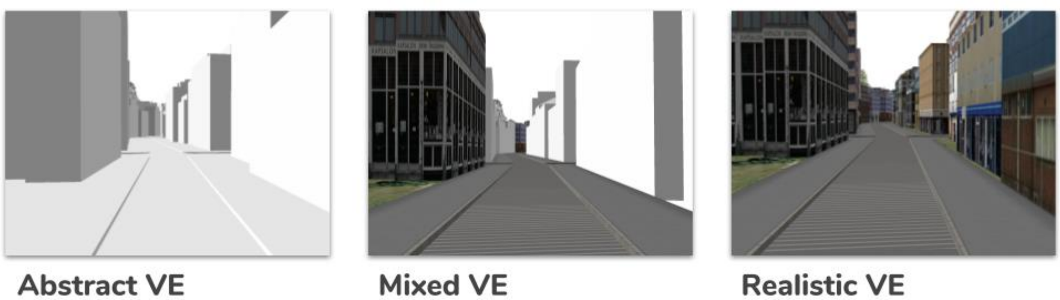


***Figure 2**: Designs of the three VEs, i.e., visualization types (recreated from Publication III).*

The research questions (RQ) in this thesis in relation to these visualization types are as follows:

Leading research question (I) How do varying levels of visual realism in 3D VEs affect the recall of visual, spatial, and visuospatial information in a virtual route learning task?

Specific research questions

RQ1.1. How does Mixed VE affect participants' recall accuracy of visual, spatial, and visuospatial information in route learning tasks when compared to the two baseline conditions (Abstract and Realistic VEs)?

RQ1.2. How do the Abstract and Realistic VEs differently affect participants' recall accuracy of visual, spatial, and visuospatial information in route learning tasks compared to each other (Abstract VE vs. Realistic VE)?

RQ1.3. How do different levels of realism affect participants' visualization type preference among the three VEs for route learning *before* and *after* the learning tasks are performed?

RQ1.4. How does the Mixed VE differently affect the alignment of participants' response confidence with their recall accuracy when compared to the two baseline conditions (Abstract and Realistic VEs)?

1.2.2 Use context: Route recall in short- and long-term

Terminology clarification: In the questions below, two *recall stages* are featured that correspond to short- and long-term memory: the *immediate* (right after the route learning experience) and *delayed* (one week later) *recall stages*. RQs related to *recall stage* are as follows:

Leading research question (II) Do the effects of varying levels of realism in 3D VEs on route recall tasks persist over time?

RQ2.1. How do different levels of visual realism affect participants' recall accuracy of visuospatial route information in the *immediate* and *delayed recall stages*?

RQ2.2. How do different levels of visual realism affect the amount of visuospatial information transferred from short- to long-term memory, i.e., what are the 'forgetting rates' in visuospatial route recall tasks from the immediate to the delayed recall stages?

RQ2.3. How does the Mixed VE differently affect the alignment of participants' response confidence with their recall accuracy when compared to the two baseline conditions (Abstract and Realistic VEs) in the *immediate* and *delayed* recall stages?

1.2.3 User characteristics: Age differences and cognitive abilities

RQs related to user characteristics in the scope of this thesis are as follows:

Leading research questions (III)

1) How do varying levels of realism in the studied VEs affect the route recall performance of healthy aging older adults and younger adults?
2) How do participants' *spatial abilities* and *visuospatial memory capacity* affect route recall with the three VEs?

Specific research questions

RQ3.1. How does the route recall accuracy of older adults differ from that of younger adults with the three VEs?

RQ3.2. How do different VEs affect younger and older participants differently in self-assessing how well they performed in the given tasks?

RQ3.3. How do participants' visualization type preference differ based on age for the three VEs for route learning tasks *before* and *after* the VE experience?

RQ3.4. How do participants' spatial abilities as measured by mental rotation task (MRT), and visuospatial memory capacity as measured by the visuospatial memory capacity test (VSM) interact with their route recall accuracy with the use of the three VEs taking also age into account?

1.2.4 Structure of the thesis

Chapter 1 provides the overall motivation for this Ph.D. project, presents the RQs, provides an overview of the dissertation's scope and content, and links the publications—on which this paper-based thesis is built—to the RQs. **Chapter 2** provides the theoretical background of the thesis, synthesizing the key literature on core topics from the sub-disciplines of geography (cartography and geovisualization, geographic information science and technology) and psychology (perception and cognition in relation to attention, memory, and aging). Thesis' aims, research gap and hypotheses are presented at the end of this chapter, linking them to the RQs. **Chapter 3** presents the methodology applied throughout the Ph.D. project, which has been used in all of the included publications. **Chapter 4** presents the main findings of the thesis. **Chapter 5** revisits the RQs and discusses the findings embedded in the relevant literature and states the limitations of the work. **Chapter 6** brings the findings together and focuses on the larger picture. To conclude the manuscript, scientific contributions of the work are revisited, and a list of visualization design guidelines are offered. Chapter 6 concludes with an outlook. **Publications** section includes a reprint of the three journal publications on which the thesis is based. **Appendix** section includes the materials used in the experiment, as well as the Ph.D. candidate's curriculum vitae.

1.3 Included publications

Throughout this Ph.D. project, I led three journal and 10 conference papers and contributed to one journal and two conference papers led by others (Figures 2, 3, and 4). This thesis is based on the three journal papers published in international peer-reviewed journals for which I am the first author. The conference papers (abstract-reviewed full papers or abstracts) are not explicitly attached to this thesis, but they are cited throughout the manuscript where appropriate.

Publication I: Lokka, I.E., & Çöltekin, A. (2019). Toward optimizing the design of virtual environments for route learning: empirically assessing the effects of changing levels of realism on memory. *International Journal of Digital Earth*, 12(2), 137-155. https://doi.org/10.1080/17538947.2017.1349842

Ph.D. candidate's contributions: Development of the concept. Implementation and execution of the experiment. Data analysis. Writing the draft of the manuscript and incorporating corrections and revisions.

Çöltekin, A., **Lokka, I.E.**, & Boér, A. (2015). The utilization of publicly available map types by non-experts—A choice experiment. In *Proceedings of the 27th International Cartographic Conference (ICC2015)*.

Lokka, I.E., & Çöltekin, A. (2016). Simulating navigation with virtual 3D geovisualizations—A focus on memory-related factors. In *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Prague. (DOI: https://doi.org/10.5194/isprs-archives-XLI-B2-671-2016)

Çöltekin, A., **Lokka, I.E.**, & Zahner, M. (2016). On the usability and usefulness of 3D (geo)visualizations—A focus on virtual reality environments. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. (DOI: https://doi.org/10.5194/isprs-archives-XLI-B2-387-2016)

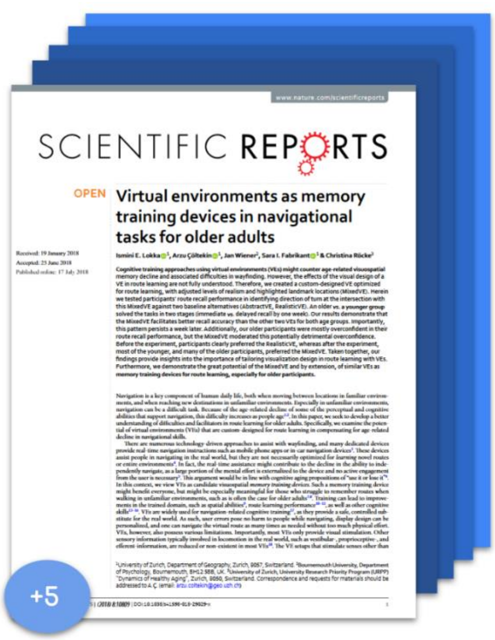
Lokka, I.E., Çöltekin, A. (2017). Remembering what we see: Designing virtual environments to improve visuospatial recall for navigation tasks. In *Proceedings of the 28th International Cartographic Conference (ICC2017)*.

Lokka, I. E., & Çöltekin, A. (2018). A virtual reality experiment for improving the navigational recall: What can we learn from eye movements of high-and low-performing individuals?. In *Proceedings of the 3rd International Workshop Eye Tracking for Spatial Research (ET4S)*. ETH Zurich. (DOI: https://doi.org/10.3929/ethz-b-000222473)

***Figure 3**: Introductory page of **Publication I** (left), and the thematically related conference publications (right).*

Publication II: Lokka, I.E., Çöltekin, A., Wiener, J. M., Fabrikant, S. I., & Röcke, C. (2018). Virtual environments as memory training devices in navigational tasks for older adults. *Scientific Reports*, 8(1). https://doi.org/10.1038/s41598-018-29029-x

Ph.D. candidate's contributions: Development of the concept. Implementation and execution of the experiment. Data analysis. Writing the draft of the manuscript and incorporating the contributions from all co-authors, as well as corrections and revisions.



Lokka, I.E., Çöltekin, A. (2017). Virtual environments as memory training devices for navigational tasks as we age: A design perspective. *In 2nd International Workshop on Models and Representation in Spatial Cognition*, April 6-7, Tuebingen, DE.

Lokka, I.E., & Çöltekin, A. (2017). Designing memorable 3D geovisualizations for older adults. *Aging & Cognition 2017,* EUCAS, April 20-22, Zurich, CH.

Lokka, I.E., & Çöltekin, A. (2017). Navigational learning in virtual environments that are designed to improve memory—Individual and group differences based on spatial abilities and age. Workshop on *Urban Wayfinding & the Brain*, June 14th, 2017, UCL, London, UK.

Lokka, I. E., & Çöltekin, A. (2018). Virtual environments designed to improve route learning performance: A focus on age and visuospatial abilities. *ICA Commissions Joint Workshop Altases, Cognition, Usability, April 2018.* Olomouc, Czech Republic.

Lokka, I.E., & Çöltekin, A. (2019). Age differences in attention and memory in a virtual reality route learning task. In *5th International Conference Aging & Cognition* EUCAS2019. Apr 24-26.

***Figure 4**: Introductory page of Publication II (left), and the thematically related conference publications (right).*

Publication III: Lokka, I.E., & Çöltekin, A., (2020). Perspective switch and spatial knowledge acquisition: effects of age, mental rotation ability and visuospatial memory capacity on route learning in virtual environments with different levels of realism. *Cartography and Geographic Information Science*, 47(1), 14-27. https://doi.org/10.1080/15230406.2019.1595151

Ph.D. candidate's contributions: Development of the concept. Implementation and execution of the experiment. Data analysis. Writing the draft of the manuscript and incorporating corrections and revisions.


Cartography and Geographic Information Science

ISSN: 1523-0406 (Print) 1545-0465 (Online) Journal homepage: https://www.tandfonline.com/loi/tcag20

Perspective switch and spatial knowledge acquisition: effects of age, mental rotation ability and visuospatial memory capacity on route learning in virtual environments with different levels of realism

Ismini E. Lokka & Arzu Çöltekin

To cite this article: Ismini E. Lokka & Arzu Çöltekin (2019): Perspective switch and spatial knowledge acquisition: effects of age, mental rotation ability and visuospatial memory capacity on route learning in virtual environments with different levels of realism, Cartography and Geographic Information Science, DOI: 10.1080/15230406.2019.1595151

To link to this article: https://doi.org/10.1080/15230406.2019.1595151

Published online: 23 Apr 2019.

Submit your article to this journal

Article views: 106

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=tcag20

+2

Lokka, I. E., & Çöltekin, A. (2018). Do age differences affect performance in 2D sketching based on a first-person perspective (3D) route learning task in differently-designed virtual environments? *Spatial Cognition 2018.* Tuebingen, Germany, September 05-08, 2018.

Lokka, I. E., & Çöltekin, A. (2018). Evaluating route learning performance of older and younger adults in differently-designed virtual environments: A task-differential analysis. In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-4, 383-387. Delft. Netherlands. October 1st-5th, 2018. (DOI: https://doi.org/10.5194/isprs-archives-XLII-4-383-2018)

Figure 5*: Introductory page of* ***Publication III****, and the thematically related conference publications (right).*

1.4 Linking publications with research questions

The links between the RQs and the publications are clarified below.

1.4.1 RQ 1: Visualization type: Visual realism

The key contributions of this thesis are on the effects of visual realism in VEs on navigational route learning, with a specific focus on an older age group. The three papers included in the dissertation thus reflect this focus. **In Publication I** (Lokka & Çöltekin, 2019), the broader focus is on the improvement of the visualization design of VEs for navigational tasks. By performing an extensive investigation across multiple tasks, three visually distinct VEs varying in levels of visual realism were created and evaluated in order to understand how well young adults could recall visuospatial information (n = 42). This paper essentially benchmarks the main idea of information overload and motivates the next steps. **In Publication II** (Lokka, Çöltekin, Wiener, Fabrikant, & Röcke, 2018), the focus is on age differences in relation to visualization type. The three VEs are comparatively assessed to examine how well they can assist younger *and* older adults in the recall of visuospatial information during route learning. In addition to participants' recall accuracy, their perceived accuracy (confidence) and preferences among the VEs are also analyzed (n = 81). **In Publication III** (Lokka & Çöltekin, 2020), a complex task that requires perspective switching between first-person

3D and areal 2D views is examined in connection with the visualization type; analyzing the interactions between four factors (*age*, *visualization type*, *abilities,* and *recall stage*) as independent variables (n = 81).

1.4.2 RQ 2: Use context: Recall stage

Since this project is broadly about spatial knowledge acquisition during route learning, the question of *how long* the acquired knowledge is retained is a key question. In other spatial acquisition studies, long-term recall appears to be scarcely evaluated, possibly due to its significant implementation complexities. In this project, two *recall stage*s are studied: *Immediate recall* refers to participants' recall rates immediately after experiencing the three VEs, while *delayed recall* refers to the same one week later. **Publication I** presents an extensive evaluation of participants' visual, spatial, and visuospatial information recall accuracy during the *immediate* and *delayed recall stages* (elaborated in Section 3.6) comparatively for the three VEs. **Publication II** and **Publication III** similarly evaluate the effect of *recall stage* with the three VEs in the specific recall tasks they examine (*turn-by-turn* recall and *map-sketching,* respectively). The key findings are summarized and interpreted in Chapters 4 and 5 of this manuscript and elaborated further in each published paper.

1.4.3 RQ 3: User characteristics: Age and cognitive abilities

Since any learning and recall task highly depends on memory capacity, which declines over the life course, the effects of aging are equally important in this thesis. Specifically, the effects of aging on route learning and the related visual, spatial, and visuospatial information recall tasks, as well as how these interact with the tested visualizations are treated and investigated as a third objective. **Publication II** and **Publication III** discuss the importance of age as a factor in route learning tasks. These two papers elaborate the specific challenges that especially older people face in route recall tasks of visuospatial information, for turn-by-turn recall and map-sketching. Thus, they present the differences in younger and older participants' recall performance with the different *visualization types*. Age appears as an independent variable in both papers with a detailed discussion on the differences between the two age groups that took part in the study.

Besides the age-related cognitive decline, in navigation-related literature, it is evident that participants' memory capacity and spatial abilities can interfere with route recall measures (Wolbers & Hegarty, 2010). In **Publication I**, the variability in recall performance is better explained with the introduction of group differences based on spatial abilities and visuospatial memory capacity in all the examined tasks and both the *recall stages*. Similarly, **Publication III** provides evidence that the variability of recall performance for the map-sketching task that requires a change of perspective is explained at least partially by differences in memory capacity and spatial abilities. In both papers, these two abilities are analyzed as moderating factors.

2. Theoretical background

This thesis is embedded in geographic information science (GIScience), bridging interdisciplinary knowledge from geography, visualization, and perceptual and cognitive psychology. Hence, this chapter first summarizes the overall role of visualizations in science, then narrows it down to 3D visualizations used in geography, and further focuses on the role of visual realism in visuospatial knowledge acquisition in the context of virtual navigation and route learning. Then the question of "for whom" is examined: The role of memory capacity in visuospatial recall, as well as visual and spatial memory types are reviewed as cognitive factors in route learning. Finally, the literature on age-related cognitive decline in healthy aging individuals in the context of navigation is covered. All these factors are then brought together, establishing the basis of the thesis. After the review of the literature, the thesis aims and research gaps are presented.

2.1 3D and visual realism in visualizations

This thesis was initially motivated by the fact that the value of 3D and realism as a *visualization type* has been subject to scholarly debate (Boér, Çöltekin, & Clarke, 2013; Çöltekin, Lokka, & Zahner, 2016; Shepherd, 2008a; Smallman & John, 2005b; Wood, Kirschenbauer, Döllner, Lopes, & Bodum, 2005). Examining the key literature from this perspective will set the foundation and guide the thinking behind how one should design realistic 3D visualization environments that facilitate effective route recall in navigational VR experiments. Such an investigation could lead to visualization design guidelines, creating "cognitive amplifiers" for younger people and serving as "cognitive prosthetics" for older people (Arias-Hernandez, Green, & Fisher, 2012; Card, Mackinlay, & Shneiderman, 1999).

2.1.1 2D and 3D visualizations in science

To examine the value of 3D visualizations, first we examine the fundamental concepts in 2D visualization literature. Bertin's seminal "visual variables" are considered the building blocks of a (2D) visualization (Bertin, 1983). Modern scholars extend Bertin's visual variables to interactive visualizations (Carpendale, 2003; Dibiase, Maceachren, Krygier, & Reeves, 1992; MacEachren, 2004), as well as 3D (Slocum, McMaster, Kessler, & Howard, 1999) by adding variables such as dimensionality, interaction modality, animation, abstraction, shading, camera angles (perspectives) in 3D, etc. Many of these factors have been examined in user experiments, albeit to a limited degree, where researchers aim to provide guidelines for their use and design (Boukhelifa, Bezerianos, Isenberg, & Fekete, 2012; Brügger, Fabrikant, & Çöltekin, 2017; Brychtova & Çöltekin, 2016; Garlandini & Fabrikant, 2009; Maggi, Fabrikant, Imbert, & Hurter, 2016). The use of 2D or 3D displays, with the question of "when is one more appropriate than the other?", appears to have attracted some interest throughout the past decades (Cockburn & McKenzie, 2002; Risden, Czerwinski, Munzner, & Cook, 2000; Tory, 2003). 2D visualizations are well-established means of visualizing geospatial content. The fields of map-making and cartography have a long history of examining what constitutes a good map. However, the technological developments since the 1960s truly accelerated through the 1980s to today, enabling computer graphics communities to bring the third spatial dimension as a possibility for representing information. This "extra dimension" was viewed as an opportunity to present more information (Dykes, MacEachren, & Kraak, 2005). While the value of a 2D visualization has been well-established for various use cases (e.g., presentation of graphs, cadaster maps, etc.), the use of the third dimension seemed random in general, lacked evidence regarding whether it supported user performance with displays, and attracted criticism (Shepherd, 2008; Tufte, 2001). Information visualization literature suggests that a justified use of the third dimension in visualizations is yet to be established (e.g., Borkin et al., 2011; Çöltekin et al., 2016), whereas in scientific visualization communities 3D visualizations appear to be more accepted (Çöltekin, Bleisch, Andrienko, & Dykes, 2017).

In the geovisualization community and related areas, those who study visualizations in connection to human cognition have been investigating the use of 2D and 3D geovisualizations in comparison to one another as well; for example, examining the combinations of the two (Bleisch & Nebiker, 2008; Herbert & Chen, 2015; Seipel, 2013), or comparing stereoscopic 3D with monoscopic 3D (Fabrikant, Maggi, & Montello, 2014; McIntire, Ellis, Harrington, & Havig, 2014; Shepherd, 2008). As in other visualization communities, the value of the additional spatial dimension is debated in geovisualization, as studies that investigate how accurately people can understand geospatial content and make decisions using 2D and 3D maps demonstrate mixed outcomes (Borkin et al., 2011; Çöltekin et al., 2016; Huk, 2006; Zanola, Fabrikant, & Çöltekin, 2009). The reason for such mixed findings in user studies with 3D in visualizations is likely due to the varying tasks (use cases and context) and participant characteristics, though clearly more work is needed to establish the exact reasons behind the different findings. In navigation-related research, we see a mixture of 2D and 3D geovisualization use (Meijer, Geudeke, & Van Den Broek, 2009). Similarly, navigation assistants such as Google Maps typically use abstract 2D representations on their landing page, whereas for turn-by-turn directions during navigation, we see a shift to 3D first-person views (Google Maps[3]). This shift can be explained due to the change in task. On the landing page, the 2D aerial view ensures that the user sees the route in its entirety and can get an overview of the spatial context, as well as their absolute and relative orientation (Thorndyke & Hayes-Roth, 1982). With turn-by-turn directions in real-world navigation, a 3D first-person perspective ensures that the contents of a user's viewpoint directly links to their real-world experience; thus, no mental rotations are required to bring the scene into the same perspective (Golledge, Dougherty, & Bell, 1995). Importantly, during real-world navigation, there is evidence that people appreciate direct links between the visualization and the real world (Burigat & Chittaro, 2007; Liao, Dong, Peng, & Liu, 2017; Plesa & Cartwright, 2008). On the other hand, it has also been documented that people are not always able to judge if what they prefer is what better facilitates their performance (Smallman & John, 2005a). In the scope of this thesis, a direct comparison of 2D vs. 3D is not made. However, the perspective switch is examined, i.e., if the spatial knowledge acquired in the 3D first-person view VEs could be successfully translated to 2D sketches by participants (**Publications I, and III**).

2.1.2 Visual realism

3D visualizations can simulate the real world with high fidelity. Thus, a common argument for their use is that they are more "natural" to humans, as the world they experience is 3D (Finlayson et al.,

[3] https://google.com/maps

2017). Assuming this is overall a good thing, a decade-old question linked to this argument is "How high-fidelity must a representation be for it to look and feel real?" (Luebke et al., 2003). The fidelity of the representation can be too low, which would not feel real, or it can induce the "uncanny valley" effect wherein a sense of familiarity is achieved but is still unconvincing, or the visualization evokes negative emotions (MacDorman, 2006; Mori, MacDorman, & Kageki, 2012). Of course, the level of realism in a visualization can also be "just right". In cartography and geovisualization discourse, this set of concepts regarding realism levels in representation is studied as *levels of abstraction/realism* and is a long-standing challenge (Bétrancourt & Tversky, 2000; Çöltekin, Bleisch, Andrienko, & Dykes, 2017; Cover et al., 1993; Gonzalez, 1996). This is similar to the debate on whether one should or should not use 3D: In visualization-related literature, as well as in perceptual psychology studies, the evidence for and against the use of visual realism is mixed (McIntire, Havig, & Geiselman, 2014; Smallman & John, 2005b). One can find both support and opposition for the use of visual realism to communicate information or to acquire spatial knowledge from visual displays.

A number of studies support the position that with the removal of irrelevant information from the visualization, user performance in various spatial tasks would increase (Çöltekin et al., 2018; Krejtz, Çöltekin, Duchowski, & Niedzielska, 2017; Smallman & John, 2005a; Wilkening & Fabrikant, 2013). Previous work suggests that when *visual clutter* on a display decreases (Rosenholtz, Li, & Nakano, 2007), the cognitive load for processing the information also decreases (Sweller, 1988). Therefore, with abstract displays, viewers have a better chance of focusing on task-specific information in achieving their goals. One of the general principles of cartography is based on this assumption: Depending on the cartographic scale, map purpose, and audience, cartographers apply *generalization* techniques to remove redundant information, highlight relevant objects, and bring a semantic focus to the map to help people see important information (McMaster & Shea, 1992).

In studies that examine whether realism adds value in (geo)visualizations compared to their abstract alternatives, it is demonstrated that users' preference and performance in given tasks (e.g., task accuracy, completion time) do not always coincide. Smallman and John (2005a) refer to this mismatch as "naïve realism": People prefer displays with a high degree of visual realism but they perform worse with realistic displays compared to abstract alternatives (Smallman & John, 2005b). This idea conceptually extends to cartography as *naïve cartography* (Hegarty, Smallman, Stull, & Canham, 2009): In a study featuring meteorological maps, non-expert participants preferred the 3D and animated displays but they performed faster with the more abstract looking displays. Later studies broadly confirm the presence of the naïve cartography effect (*e.g.,* Brügger et al., 2017). While realism may not always support performance in, for example, map reading, there is also evidence in favor of using visual realism in some contexts. For example, it has been shown that people are good at interpreting pictures and images of the natural world and understanding the gist of a scene quickly (Wolfe, 1998). It has also been shown that humans are excellent at recalling pictures that they have been exposed to in as little as a few seconds (Potter & Levy, 1969). When it comes to the amount of visual information, participants recognized thousands of images (around 2.5k), depicting even elements from the same themes, with very high accuracy (87–92%) (Brady, Konkle, Alvarez, & Oliva, 2008). Importantly, the *human recognizable elements* in realistic displays assist users in achieving higher recall rates (Borkin et al., 2013; Brady et al., 2008; Isola, Xiao, Torralba, & Oliva, 2011) as people can easily name them. The "nameability" of features appears to be important in recall tasks: It has been shown that not only objects but also colors that people can name are better recalled (Brewer, 1996; Brown, Lindsey, & Guckes, 2011; Lenneberg, 1961; Özgen, 2004). Possibly motivated by such studies, there have been efforts to increase realism levels in VR in navigation-related domains (Betrancourt, 2005; Peters, Jahnke, Murphy, Meng, & Abdul-Rahman, 2017; Semmo, Trapp, Kyprianidis, & Döllner, 2012) so that people can easily match real-world structures and landmarks to the ones on the displays (Partala et al., 2010).

Based on the above summary of the literature on realism and abstraction, one may also ask "How much (realism) is too much?" A seminal study in psychology suggests that human working memory can only store up to seven items, give or take two (Miller, 1956). Cowan's more recent findings suggest that the "magical number" might be four rather than seven (Cowan, 2010). The use of photorealism in representing four or seven items (or categories) is not straightforward. However, these papers highlight the limits of human working memory and the necessity for the adaptation of visualizations to match human limitations in processing information. In this thesis, levels of visual realism in designing VEs for navigational route learning is a core interest. Some links between visual realism and information recall were introduced above, and the following sections further elaborate on the current state of the art on human memory and information processing.

2.2 The role of memory capacity in navigation

Because the *context* in this thesis is navigational route learning, and route learning requires the use of various human memory systems, this section summarizes how memory works in a broad sense, with a focus on discourse relevant to the thesis.

What exactly is memory? Memory is a difficult-to-define term with many layers of complexity. It has been the subject of debates in the literature. Fundamentally, it is understood that it consists of psychophysical structures and processes required to store and retrieve information (McLeod, 2017). According to a commonly cited cognitive model, the three basic processes for memory are the *encoding*, *storage*, and *retrieval* of information (Melton, 1963). In order to create a new memory, the information first has to be encoded, creating a path in the brain, storing the information, and linking the specific event to the retrieval process for later access. Encoding and storage usually happen passively while retrieval is often active (Mastin, 2010). Throughout this manuscript, the word *recall* is used to express the processes involved in *information retrieval from memory*.

When considering memory, time is important. The more time that passes, the fewer details we can recall. Initially, when there is a stimulus, our senses detect it, and either ignore it, after which the stimulus disappears, or perceive it, at which point the stimulus is stored in *sensory memory* (Dubuc, 2002). At the sensory processing stage, attention is not required, but once attention is focused on an object and it acquires a meaning, this stimulus is then stored in *short-term memory* (Ware, 2004). The capacity of short-term memory is limited, and the duration that a stimulus can be retained in it is less than a minute (Terry, 2016). *Working memory* (WM) is closely related to short-term memory, and sometimes the two terms are used interchangeably. WM also has a limited capacity. It can hold four-to-nine items simultaneously (depending on the study) (Miller, 1956; Cowan, 2010). With active effort such as repetition or combining the stimulus with previous knowledge and providing it with meaning, information can be transferred to *long-term memory*. This is also the final step at which point the memory can be stored for an extended amount of time, some even for an entire lifetime. The capacity of long-term memory is "unlimited" (McLeod, 2010).

2.2.1 Memory and learning for routes

Visual and spatial memory Apart from the temporal considerations in discriminating between memory types (short-term or working memory, long-term memory), most memory models suggest that the encoding of stimuli can occur in different forms and in specialized parts of the brain. This section includes a concise review of *visual* and *spatial* memory types, both of which are required during navigation, given that navigation includes a combination of visual and spatial components. The distinction between the visual and spatial memory types is somewhat fuzzy, though some tasks show clear dissociation between visual and spatial components *e.g.,* as demonstrated in the Visual Patterns Test and the Corsi Blocks Test (Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999). Distinguishing the terms *visual* and *spatial,* Logie (1986) states: "One way in which to think of the term 'spatial' is as a reference to the location of items in space and the geometric relationships between those items. Visual information might then refer to properties of those items such as their shape, color, and brightness" (pp. 77–78). When a task requires both memory types in tandem, which is typical of many of the tasks linked to route learning, the term *visuospatial memory* is used. Although not treated in this project, another interesting memory system worth mentioning is the *episodic memory* (Tulving, 1972) which may be relevant to route learning (Gras, Gyselinck, Perrussel, Orriols, & Piolino, 2013). Episodic memory is distinguished from other forms of memory, as the experiences are mentally framed in a certain time and location, and can be recalled together as a set (Tulving, 1972). Information from multiple channels (visual, spatial, verbal) might be linked in episodic memory (Lee & Spelke, 2010). Route recall requires the use of all these memory types, and environmental cues play a crucial role in the type of information encoded and recalled.

The importance of landmarks As explained in Section 2.1, there is a link between the *amount* of information presented to people and *how much* they can remember. Besides the *amount* or *quantity* of the information, various semantic and perceptual features, i.e., its *quality,* need to be taken into consideration (Lokka & Çöltekin, 2017). In navigation, landmarks have such perceptual and semantic qualities that are highly relevant for spatial knowledge acquisition. Overall, landmarks can be described as salient objects or features in the environment that attract attention, and thus, help people in creating mental anchors, or spatial frames of reference, linked to a specific location which is important in the memorization of a route (Richter & Winter, 2014). How well people describe a route is shown to be correlated with the number of landmarks they can recall (Heft, 1979).

It is important to note that there are different types of landmarks. The most common categorization of landmark types is based on *why* they may attract attention. Thus, a landmark's i) visual, ii) semantic, and iii) structural salience and a combination of these factors appear to determine their level of *landmarkness* (Aginsky, Harris, Rensink, & Beusmans, 1997; Klippel & Winter, 2005; Raubal & Winter, 2002; Sorrows & Hirtle, 1999; Strickrodt, O'Malley, & Wiener, 2015; Waller & Greenauer, 2007; Wiener, de Condappa, Harris, & Wolbers, 2013). The *visual* salience of a landmark relates to its distinguishable features that make it stand out from the rest of the objects in its surroundings, *e.g.,* shape, size, color, texture, etc. *Semantic* salience relates to the general or personal meaning of an object or location *e.g.,* temple, school, hospital, etc., or a location wherein a person may have experienced a memorable moment. *Structural* salience relates to the location of the object in relation to its immediate environment e.g., on vs. off a route of interest, at intersecting points of the path, etc., and it is essentially tied to the characteristics of the route of interest (Claramunt & Winter, 2007; Klippel, Richter, & Hansen, 2005; Klippel & Winter, 2005; Lenneberg, 1961; Röser, Hamburger, Krumnack, & Knauff, 2012; Waller & Lippa, 2007). In another theoretical study, Stankiewicz & Kalia (2007) approach landmark characterization based on: (i) *persistence*, i.e., how stable a landmark is in position, (ii) *perceptual salience* (same as above), and (iii) *informativeness* for the navigational task.

Furthermore, Waller & Lippa (2007) distinguish landmarks based on whether they serve as *associative cues for action*, or as *beacons*. If landmarks serve as associative cues for action, "recognition triggered responses" are introduced, i.e., people are asked to take an action as soon as they identify a landmark. The call to action in such cases can be disconnected from the overall navigational goal of arriving at the destination, and focused on a specific segment of the route, e.g., taking a left turn at the gas station. The term associative cue is linked to '*cued recall*', in which learning occurs by associating a specific action (turning left), with the specific landmark (gas station). In this process, one needs to recall and preserve two types of information: landmark identity, and directional information. Therefore, the time required to memorize the route based on such landmarks may be longer (Waller & Lippa, 2007). On the other hand, the concept of "landmarks as beacons" is relevant when one is aiming to reach the identified landmark, which may be their final goal or allow them to finish a segment of the route. Beacons are linked to *recognition* tasks, where a person recognizes a specific landmark that leads her to the final target. In this case directional information is not necessarily present, but the correct identification of the next beacon is important, with a route resulting in a combination of many successive landmark beacons (Waller & Lippa, 2007).

Among the various landmark features discussed above, the structural salience is perhaps specifically interesting from the perspective of spatial sciences. There are numerous navigation studies with the goal of identifying optimal locations for structurally salient landmarks (Karimpur, Röser, & Hamburger, 2016; Richter & Winter, 2014; Röser, Hamburger, et al., 2012; Röser, Krumnack, Hamburger, & Knauff, 2012; Winter, 2003). The outcomes from these studies suggest that landmarks at specific locations on a person's route in wayfinding tasks signify high structural salience. Specifically, intersections (decision points for a turn) and landmarks located in the direction of the turn are critically important (Röser, Hamburger, et al., 2012). These are the positions on a route where people take "mental notes" of what is around them (named "effective landmark locations" in this thesis) to recall their route later. In addition to these landmarks, the spatial layout of the street network (structural network) serves as another anchor that people utilize to learn their route while navigating in unfamiliar environments (Claramunt & Winter, 2007).

Viewing perspective and route learning strategies The perspective under which the learning occurs can have a profound effect on one's acquiring spatial knowledge. In the real world or in a VE, an observer typically experiences the environment from a first-person perspective, i.e., from an *egocentric* view (Klatzky, 1998). Learning can also occur from a bird's eye view—for example, using a map—where the environment is presented in a reference frame that is external to the observer, i.e., from an *allocentric* view (Klatzky, 1998). Both these perspectives are important in spatial knowledge acquisition and allow building *route* and *survey* knowledge (Montello et al., 2004).

2.3 The role of age and cognitive abilities in route learning

There are large individual and group differences in people's navigation abilities, as well as their abilities to read, interpret and recall visualizations (Wolbers & Hegarty, 2010), all of which are important in the context of this thesis. A considerable variation in navigation and spatial task performance has been documented due to individual and group differences, although these experiments measure spatial learning in the short term (Brügger, Richter, & Fabrikant, 2019; Credé, Thrash, Hölscher, & Fabrikant, 2020; Meneghetti, Borella, Muffato, Pazzaglia, & De Beni, 2014). An individual's prior experience, cognitive abilities, or age are just a few in a long list of well-known factors that may affect performance in spatial tasks. In this Ph.D. thesis, as mentioned in earlier sections, age is of particular interest due to its impact on memory, whereas several other factors such as gender and expertise have been controlled (see Chapter 3). Below, a concise review of the literature is provided to further justify why *age* is an important and interesting factor to study in connection to navigation-related spatial tasks that require remembering.

2.3.1 Memory throughout the lifespan

Changes in cognition in healthy aging Cognitive psychology literature clearly demonstrates that age is one of the most significant factors affecting cognitive abilities, including memory capacity (Anders, Fozard, & Lillyquist, 1972; Foos, 1989; Park et al., 2002). Numerous studies document the decline in memory capacity over the life course, e.g., in the recall of faces, text, visuals, and verbal input (Park et al., 2002; Rendell, Castel, & Craik, 2005; Salthouse, 2004). A seminal and comprehensive study on cognition throughout the lifespan (Park et al., 2002) identified that while "knowledge-based verbal abilities" increase (pp. 305), working memory, short-term memory, long-term memory, and the speed of processing exhibit a steady decline as people age (Figure 6).

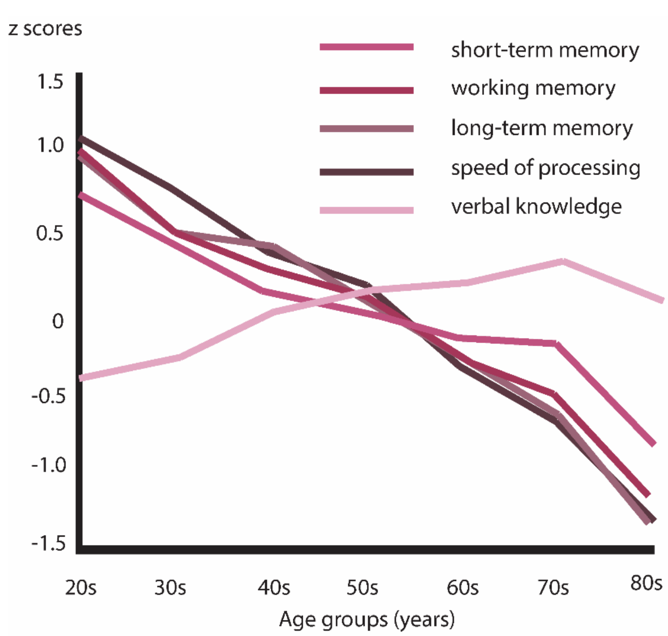


Figure 6*: A composite view of lifespan performance measures, including: speed of processing, working memory, long-term memory, short-term memory, and knowledge-based verbal abilities (redrawn from Park et al., 2002).*

In addition to the decline in memory with aging (or rather possibly because of it), older people seem to make more misattribution errors in the *location* and *time* of an experienced event (Dodson, Bawa,

& Slotnick, 2007; Dodson, Koutstaal, & Schacter, 2000; Kelley & Sahakyan, 2003; Zelinski & Light, 1988). The number of misattribution errors correlates with the number of things people have to remember (Dodson et al., 2000), confirming the cognitive load theory (Sweller, 1988). Another important issue to consider in connection with aging is that older adults tend to overestimate the amount of information they can accurately recall and as a result, they are overconfident about their memories (Dodson, Bawa, & Slotnick, 2007). Findings that support these various shortcomings have been consistently observed in verbal recall tasks e.g., memorizing lists of words (Pardilla-Delgado & Payne, 2017), as well as photo/video recall tasks e.g., in crime scene events (Dodson & Krueger, 2006; Koutstaal & Schacter, 1997). It has also been well-documented that spatial memory declines earlier than, e.g., visual or verbal memory systems (Barbeau et al., 2004).

Navigation in older age Navigation relies on multiple cognitive factors and is a complex task for most people. Age-related cognitive decline makes navigation even more complicated in older age (Klencklen, Després, & Dufour, 2012; van der Ham & Claessen, 2020), especially in unfamiliar settings (Kirasic, 1991; Moffat et al., 2001; Muffato, Della Giustina, Meneghetti, & De Beni, 2015). An age-related decline in navigation ability has been observed in both real-world and virtual reality settings, with strong links to cognitive decline which is also reflected in people's visuospatial abilities, mental rotation abilities, and memory capacity (Evans & Pezdek, 1980; Gyselinck et al., 2013; Hertzog & Rypma, 1991; Moffat & Resnick, 2002; Muffato, Meneghetti, Di Ruocco, & De Beni, 2017; Nemmi, Boccia, & Guariglia, 2017; O'Malley, Innes, & Wiener, 2018; Park et al., 2002).

Various studies compare the performance of younger and older adults in navigation related tasks. For example, healthy-aging older adults are typically slower in memorizing a novel route than young adults (Barrash, 1994; Iaria, Palermo, Committeri, & Barton, 2009; O'Malley et al., 2018). Additionally, they have more difficulty in recalling landmarks and recognizing environmental scenes (Head & Isom, 2010); locating landmarks (Gyselinck et al., 2013); recalling the route sequence (Head & Isom, 2010); and in associative route learning (Head & Isom, 2010; O'Malley et al., 2018). The evidence is consistent that aging has an overall negative influence in navigation related tasks, and most authors link this performance issue to declining cognitive abilities. Aging appears to also have a negative effect *especially* in allocentric spatial tasks in navigation (Fricke & Bock, 2018), with older adults experiencing difficulties in forming and using an allocentric cognitive map (Iaria et al., 2009; Moffat & Resnick, 2002; Wiener et al., 2013) which is possibly linked to the observation that older people commit more errors in mental rotation tasks (Herman & Coyne, 1980; Inagaki et al., 2002). These issues all affect successful navigation and related tasks, such as route learning.

2.4 Stating the gaps

The theoretical background section above highlights the most important factors in examining route learning in a VE: How do we design visual displays with a specific focus on levels of visual realism?; Why does memory capacity matter in studying route learning?; Why should participant characteristics, especially age, not be ignored in scholarly work that examines memory-related tasks? In light of the reviewed literature, the gaps in the scientific knowledge are identified, followed by the hypotheses tied to each research question presented in Section 1.2. The literature review in this chapter clearly demonstrates the progress made in various scientific disciplines on topics related to visualization, aging, and navigation. At the intersection of these domains, new questions arise on the interactions between aging, navigation, and visualization.

An important gap relates to **the effects of visual realism levels in (virtual) route learning** tasks. This is important, because with the right levels of visual realism, VEs can facilitate route learning and support people in the complex task of navigation. People seem to prefer visually realistic displays, and the VR community aspires to mimic reality in these displays. However, there is not a clear understanding of the precise effects that different levels of visual realism in VEs have on individuals' visuospatial information recall in the context of navigation. It is not yet understood what degree of visual realism effectively facilitates route recall, and simultaneously, leads to healthy levels of confidence in people regarding their own recall performance. In this thesis, the question about whether participants really prefer visually realistic displays is evaluated and the effects of varying levels of visual realism on route recall performance and confidence are investigated.

A second important gap is in the **distinction between the short- and long-term route recall.** This distinction may be the difference between memorization and learning since *learning* implies that phenomena of interest stay in the mind for more than a few minutes. Many spatial knowledge acquisition studies may not be measuring learning because they lack longer-term measurements of recall performance, likely due to the inherent practical complexities of repeating the experiment after some time has elapsed. However, if the tasks involve memory (such as route learning), these temporal considerations are important because they affect recall rates. In this thesis, a delayed recall experiment is included to better understand the longer-term effects of visual realism on route learning.

Both spatial abilities and memory capacity decline over the life course, and both play an important role in navigational recall, leading to the third important gap: The absence of an explicit examination of **how age-related cognitive decline impacts route learning with varying levels of visual realism**. Thus, relating people's age, memory capacity, and spatial abilities can shed more light on the overall route learning process. In this thesis, how different age groups perform with proposed visualizations is investigated to identify implications that visualization designs can have on route learning in relation to aging. Such investigations can further support the creation of "cognitive amplifiers" for younger people, in addition to serving as "cognitive prosthetics" for older people (Arias-Hernandez et al., 2012; Card et al., 1999). A better understanding of how varying the visual realism of VE designs might affect the route learning performance of people in different age groups and with different characteristics over the course of a shorter and longer time frame will contribute to the process of filling the above stated gaps. In broader terms, filling these gaps will advance our overall understanding of how spatial knowledge acquisition occurs during route learning.

2.5 Research hypotheses

Following the identified research gaps, I formulate literature-based hypotheses for each RQ presented in Section 1.2. These hypotheses were all tested and treated in all three publications.

2.5.1 Levels of realism

For all RQs that are directly about visualization type, hypotheses were formulated regarding the differences between the Mixed VE and the other two VEs. It is important to remind the reader that the Mixed VE retains photo-textures only in "effective landmark locations". These effective landmark locations were selected based on previous findings from navigation literature, which provides evidence that people pay attention to features and structures at the intersections during navigation, especially to those at the direction they are about to turn (Röser, Hamburger, et al., 2012), and based on the cognitive load theory that posits human information processing capacity is limited (Miller, 1956). Placing the photo-textures selectively, and only in navigation-relevant location "balances" the cognitive load (Sweller, 1988), and removes visual clutter (Rosenholtz et al., 2007), resulting in a selective highlighting using photo-textures. Using photorealism for highlighting (rather than e.g., color) is due to its "native" to virtual reality representations, and because it contains human recognizable and nameable elements which are known to support recall (Borkin et al., 2013; Brady et al., 2008; Isola, Xiao, Parikh, Torralba, & Oliva, 2014). The literature cited in this paragraph remains relevant to all hypotheses on visualization type, and is not repeated anymore for clarity and conciseness. When hypotheses require a citation not mentioned in this paragraph, relevant literature is explicitly cited next to the hypotheses.

H-RQ1.1. Based on the literature cited above, three hypotheses were developed in relation to RQ1.1.:

A. **The Mixed VE will be superior to the other two VEs in facilitating participants' route recall accuracy across visual and visuospatial recall tasks that require the recall of some visual cues, whereas for spatial tasks, photorealism should not offer an advantage**: Thus, irrespective of their age, participants' recall performance across visual and visuospatial recall tasks will be best with the Mixed VE.
B. **The Mixed VE will be superior to the other two VEs in facilitating participants' route recall accuracy in perspective-preserving visuospatial tasks:** As the critical visuospatial information for route recall (Röser, Hamburger, et al., 2012) remains present and highlighted in the Mixed VE during both the encoding and decoding of information, it will help participants, irrespective of their age, to identify the direction of the turn at intersections better than the other VEs.
C. **The Mixed VE will be superior to the other two VEs in facilitating participants' route recall accuracy in perspective-switching visuospatial tasks**: The highlighted elements in the 2D basemap of the Mixed VE will serve as anchoring points to assist people, irrespective of age, to perform the perspective switch (Thorndyke & Hayes-Roth, 1982).

H-RQ1.2. Participants' *overall recall performance* with the Realistic VE will be better than with the Abstract VE: As the Realistic VE provides more visual cues that include human recognizable and nameable elements known to support recall (Borkin et al., 2013; Brady et al., 2008; Isola et al., 2014).

H-RQ1.3. Participants will prefer the Realistic VE before the experiment: All (both older and younger) participants will prefer the Realistic VE to the other two VEs before the experiment (Smallman & John, 2005b; Smallman & Cook, 2011).

H-RQ1.4. The Mixed VE will be superior to the other two VEs in facilitating the *alignment of confidence with recall accuracy.* Balanced information and reduced visual clutter (Rosenholtz et al., 2007) will assist participants in identifying places they can and cannot recall, better aligning their confidence with their performance.

2.5.2 Short- and long-term route recall

For each RQ related to short- and long-term recall, hypotheses were developed based on the relevant scientific literature and are presented below:

H-RQ2.1. All participants' recall performance will be best with the Mixed VE, also in the delayed recall stage: Due to the fact that these human recognizable elements assist recall also in the long-term (Brady et al., 2008).

H-RQ2.2. The differences from the immediate to the delayed recall stage will be smaller with the Mixed VE: The combination of the balanced cognitive load (Sweller, 1988) and the presence of nameable and memorable realistic elements (Borkin et al., 2013; Isola et al., 2014; Brady et al., 2008) will support transferring higher amounts of visuospatial information to the long-term recall.

H-RQ2.3. The Mixed VE will be superior to the other two VEs in facilitating the *alignment of confidence with recall accuracy* in both recall stages: Balanced information and reduced visual clutter (Rosenholtz et al., 2007) together with the fact that human recognizable elements can persist in long-term recall (Brady et al., 2008) will assist people in identifying the places they can and cannot correctly recall, thus better aligning their confidence with their performance *also* in the delayed recall stage.

2.5.3 Age differences and cognitive abilities in navigation

Similar to the previous sections, for each RQ, following hypotheses were developed based on the relevant scientific literature:

H-RQ3.1. Age and recall: Younger participants' route recall accuracy will be higher than older participants' irrespective of the VE type (Fricke & Bock, 2018; Moffat et al., 2001; Muffato et al., 2015).

H-RQ3.2. Age and confidence: Overall, older participants will be overconfident in their responses in comparison to younger participants (Dodson, Bawa, & Krueger, 2007). The Mixed VE should moderate this effect (similar to H-RQ1.4) and the effect should be more pronounced for the older group.

H-RQ3.3. Age and VE preference: *Before* the experiment, all participants will prefer the Realistic VE. *After* the experiment, more of the younger participants (compared to the older) will change their preference for the Mixed VE. Older participants, however, due to an overall decline in their spatial abilities (Moffat & Resnick, 2002), might not be able to identify which visualization supports them better and, thus, will still prefer the Realistic VE after the experiment (Smallman & John, 2005b; Smallman & Cook, 2011).

H-RQ3.4.

A. **Spatial abilities and memory capacity:** Participants with higher Mental Rotation Test (MRT) or Visuospatial Memory Test (VSM) scores will outperform the participants with lower MRT/VSM scores, irrespective of age or VE type (Muffato et al., 2017; Wolbers & Hegarty, 2010).
B. **Memory capacity and visual cues:** Irrespective of age, participants with a higher VSM will outperform the participants with a lower VSM in producing accurate sketches, particularly with the Mixed and Realistic VE, as these provide potentially helpful photorealistic cues (see H-RQ1.1.C) (Vandenberg & Kuse, 1978), whereas MRT will be most relevant to the Abstract VE because this visualization type contains no photorealistic cues (Ekstrom, French, Harman, & Dermen, 1976).

3. Methodology

To test the hypotheses presented in Section 2.5 and answer the RQs presented in Section 1.2, a large controlled laboratory experiment with multiple sub-components was conducted (Figure 7).

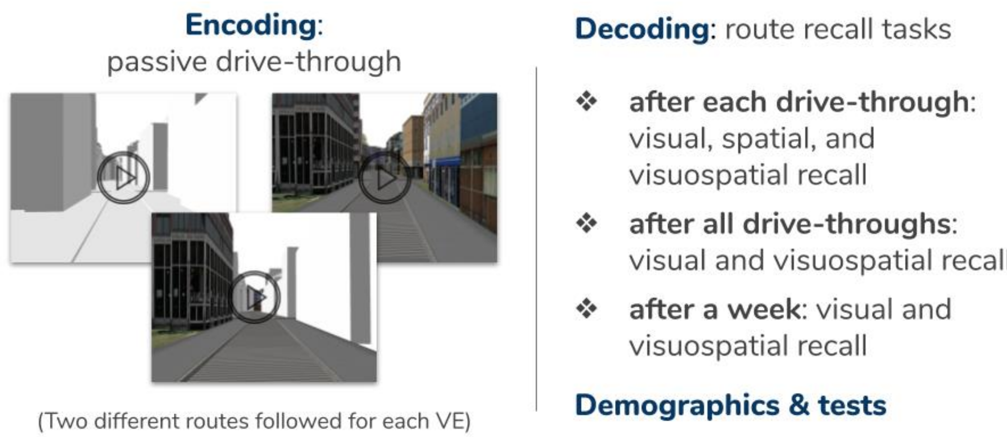


***Figure 7**: An overview of the experimental process of the thesis.*

For coherence, all experimental variables were considered together in one single large experiment design. Using a mixed factorial design, the experimental sessions were set up to investigate the interactions between specific route recall-related *tasks* and the three *visualization types,* with varying levels of visual realism while testing different *groups* of people (younger and older participants). An overview of the tested variables can be seen in Figure 8.

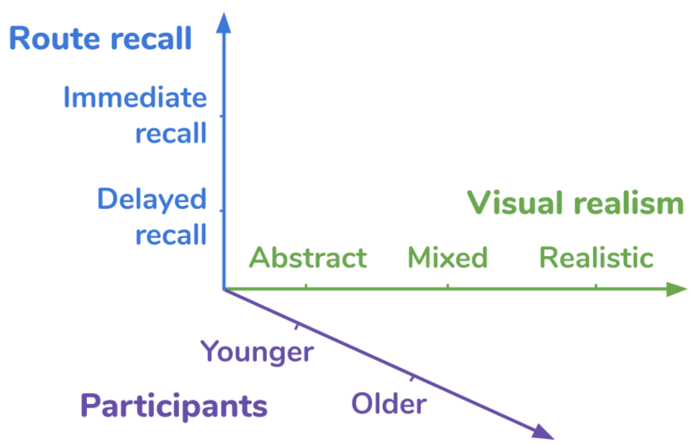


***Figure 8**: An overview of the experimental variables of the thesis.*

Section 3.6 provides further elaboration on the experimental procedure.

3.1 Independent variables

There are three main independent variables in this thesis: *age, visualization type,* and *recall stage* (Figure 8). *Age* is a binary variable that consists of younger and older people and is examined in a between-subject design. A detailed description of the two age groups follows in Section 3.4. The *visualization type*, the second main independent variable in this thesis, is designed to represent different levels of visual realism and is tested in a within-subject design. Three different VEs (Figure 2) were deliberately and carefully varied based on the photorealism they contained. A detailed description of these VEs follows in Section 3.5.2. The third and final independent variable is the *recall stage* which has a binary split (immediate and delayed) and was tested in a within-subject design.

3.2 Control variables

Some of the hypotheses presented in Section 2.5 were based on a few variables that were controlled during the experiment design, allowing for post-hoc hypotheses. These variables are 1) group differences based on spatial abilities and visuospatial memory capacity and 2) task types. For measuring participants' spatial abilities and visuospatial memory capacity, standardized psychological instruments are available and were utilized in this project, all of which are analyzed in Section 3.5.1. When the task types were designed, they were categorized into *primarily spatial,* requiring minimum photographic information, *primarily visual,* strongly requiring photographic information, and *visuospatial,* where one may have utilized photographic or non-photographic information. These three task types were not trivial to *precisely* separate; however, it was important to control for these, as the hypotheses we tested for regarding *visualization type* and *age* may have been affected by the task type. If an observed effect remains valid for all task types, it confirms that the visualization manipulation was important in all, whereas if it seems to facilitate better recall for one task type but not for the other, the conclusions require more nuance. Task types are given further elaboration in Section 3.6.1. Furthermore, participants' gender and expertise in domains related to geospatial subjects were counterbalanced; while their dominant hand, the number of languages they speak, whether they are learning a new language, their hours of sleep, their profession, their color efficiencies/deficiencies, and whether they actively train their memory were all controlled for. These variables were selected based on a comprehensive literature review and were identified as possible moderating factors (elaborated in Section 3.5.1).

3.3 Dependent variables

The main performance variable measured in the thesis was participants' *recall accuracy* of visual, spatial, and visuospatial information in the context of route learning during navigation. Recall that in navigational contexts might be based on memories encoded from direct or indirect environmental exposure, may entail differences (Montello et al., 2004). Learning an environment by physically experiencing it (by being there) is considered direct, while indirect spatial knowledge acquisition can be via photographs, maps, visualizations, etc. As this thesis' key interest is in visualizations, the measured recall accuracy is based on *indirect* information acquisition. Further variants of the recall accuracy measurements in connection with different tasks are described in Section 3.6.1. Besides recall accuracy, two key subjective variables were obtained in this thesis as dependent variables: participants' *preference* among the three VEs and their *confidence* levels in their responses. These measures complement recall accuracy and allow for building a more "holistic" understanding to inform visualization solutions for effective route learning, especially for older adults.

3.4 Participants

In total, 81 participants took part in all of the experiments. An inclusion/exclusion criterion was that the participants were *cognitively healthy* according to the Mini-Mental State Examination (see Section 3.5.1). The younger age group consisted of 42 participants (23 females) ranging from 20–30 years old, and an average age of 27. They were university students at all levels of education and of various disciplines, whereas their expertise was counterbalanced as those with expertise in spatial sciences (cartography/geography, architecture, urban planning, civil engineering, computer graphics), and those without. They were recruited by word-of-mouth. The older age group consisted of 39 participants (17 females) ranging from 65–75 years and an average age of 70. Older participants were recruited from the participant pool of the UZH's Research Priority Program, "Dynamics of Healthy Aging"[4]. The older participant sample was also counterbalanced for expertise, those with spatial sciences expertise and those without. These healthy-aging older adults are in the decade after their retirement (in Switzerland, the retirement age is 65), and before the individual variability in cognitive health becomes less predictable, which is around 75 years and older (Park et al., 2002). The experiment was approved by the Ethical Committee of the Philosophical Faculty, University of Zurich, with the form "Checkliste für die Selbstbeurteilung von Studien auf ethische Unbedenklichkeit". All participants voluntarily agreed to take part in the experiment and signed a written consent form. They could withdraw their consent at any time.

3.5 Materials

3.5.1 Psychometric and self-report measures

Mini-Mental State Examination (MMSE) As unhealthy cognitive aging was beyond the scope of this thesis, the MMSE (Tombaugh & McIntyre, 1992) was used to examine if participants should be included or excluded in the experiments (Figure 9).

Mini-Mental State Examination (MMSE)

Patient's Name: ______________________ Date: __________

Instructions: Score one point for each correct response within each question or activity.

Maximum Score	Patient's Score	Questions
5		"What is the year? Season? Date? Day? Month?"
5		"Where are we now? State? County? Town/city? Hospital? Floor?"
3		The examiner names three unrelated objects clearly and slowly, then the instructor asks the patient to name all three of them. The patient's response is used for scoring. The examiner repeats them until patient learns all of them, if possible.
5		"I would like you to count backward from 100 by sevens." (93, 86, 79, 72, 65, …)

***Figure 9**: MMSE preview of the test showing the first four questions (Tombaugh & McIntyre, 1992).*

[4] https://dynage.uzh.ch/en.html

MMSE measures cognitive impairment and is normally used when screening for dementia, where the maximum score is 30. In this study, both older and younger participants who scored 27 or higher were accepted as cognitively healthy individuals (Gallagher & Keenan, 2009).

Mental Rotation Test (MRT) The MRT (Vandenberg & Kuse, 1978) measures mental rotation abilities of participants. These abilities may be important in route learning as participants' orientation and viewing perspective may change. The MRT is a popular test that presents 3D objects consisting of individual cubes: The viewer is presented with one 3D object and must identify the two that are the same as the original among four options (Figure 10). There is a time limit of six minutes for 20 questions, which is administered after a training phase. The participant is not aware of the total number of questions and is instructed to answer as fast and as accurately as possible. For each correct answer, the participant receives one point, i.e., maximum score is 40.

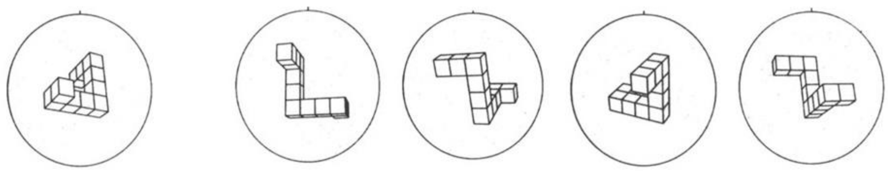


Figure 10*: Capture of the MRT. On the left is the original shape and on the right are the four options. Two represent the rotated version of the original, and two are different from the original.*

Visuospatial Memory Test (VSM) The VSM (Ekstrom et al., 1976) measures visuospatial memory capacity, which is important in route learning in general, and for tasks included in this study in particular. The VSM presents a 2D city plan with a network structure and 12 "nameable" (somewhat realistically represented) objects located in various locations all over the map (Figure 11). Viewers are given four minutes to study this plan, after which they have four minutes to locate the objects in their original position on a new sheet of the city plan without the 12 objects. After a training phase, participants do the task twice with two different city plans, each with 12 different objects on it. For each correct answer, the participant receives one point, and for each wrong answer one point gets subtracted. For each omission, zero points are awarded. The maximum score is 24.

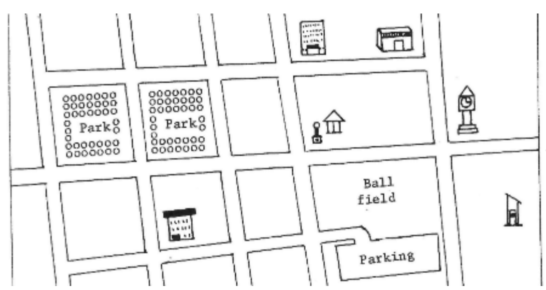


Figure 11*: A clip from one of the maps used in the Visuospatial Memory (VSM) test, in which participants had to memorize the location of the different buildings.*

Additional measures Since various other factors can impact the performance outcomes of spatial tasks using visualizations, for primarily control purposes, the following measures were additionally collected and evaluated in this thesis followed by a brief explanation of why these specific factors were selected:

- **Color blindness**: The Realistic VE uses color photo-textures, and nameable colors are more memorable for people (Brown et al., 2011; Lenneberg, 1961; Özgen, 2004).
- **Handedness**: Handedness and spatial abilities may have a complex relationship (Annett, 1992; Burnett, Lane, & Dratt, 1982; Sanders, Wilson, & Vandenberg, 1982).
- **Gender:** Depending on the test, spatial abilities may interact with gender (Coluccia & Louse, 2004; Grön, Wunderlich, Spitzer, Tomczak, & Riepe, 2000; Wolbers & Hegarty, 2010).
- **Hours slept:** Lack of sleep negatively affects memory (Maquet, 2001).
- **Expertise** with visualization and spatial sciences (cartography, geography, 3D visualizations, 3D city models, graphic design, and photography): Expertise has been consistently shown to affect success in spatial tasks (Maguire, Woollett, & Spiers, 2006).
- **Educational level:** People with a higher education may have more strategies to memorize a task that is explicitly about learning (Angel, Fay, Bouazzaoui, Baudouin, & Isingrini, 2010).
- **Active memory training:** People who actively do memory training exercises may have superior strategies to others in memory-related tasks (Gross et al., 2012).
- **Number of spoken languages:** Speaking several languages may provide cognitive advantages (Bak, Nissan, Allerhand, & Deary, 2014).
- **Learning of a new language:** See above.
- **Video gaming:** Gamers might develop superior spatial skills (Schofield & Honoré, 2009).

Note that numerous correlation analyses are possible based on these variables; this is also why they are not reported in the scope of this manuscript to keep the text readable. They were analyzed where any surprising results were observed to check for biases to support the main analysis only.

3.5.2 Stimuli

The VEs were created using the ESRI City Engine[5] with procedural programming, featuring a fictitious city with buildings of similar shape, style, and color. The structural layout of the city was designed based on a mesh-like structure with intersections at 90° angles to ensure visibility from a greater distance. The Abstract VE serves as a baseline and the Realistic VE represents the "gold standard" in virtual reality (fully photorealistic) and, thus, is essentially another baseline. The Mixed VE is deliberately designed and proposed within this thesis as a visualization that would amplify cognition and memory for younger participants and would serve as a "prosthetic" for the older age group (Arias-Hernandez et al., 2012; Card et al., 1999). In the Abstract VE, some relatively subtle visual cues are available to people in greyscale rendering of a fictitious urban environment that simulate the shading provided by natural light, helping distinguish the shapes of the structures (Figure 2 left). The Realistic VE is the same urban environment designed with color photo-textures representing the real world as realistically as possible (Figure 2, right). The Mixed VE is proposed in this thesis as a customized solution that balances between a lack and an abundance of visuospatial

[5] https://www.esri.com/en-us/arcgis/products/esri-cityengine/overview

information that the other two VEs "suffer" from (Abstract VE and Realistic VE, respectively). All objects (buildings, trees, etc.) in the Mixed VE are depicted in greyscale similar to the ones in the Abstract VE, except those located in critical locations ("effective landmark locations") along the relevant route: The most critical objects are "highlighted" using realistic photo-textures (i.e., photo-textures are added—or were not removed—only in selected locations), using the same photo-textures that were used in the Realistic VE (Figure 2, middle) (Lokka & Çöltekin, 2017). Additional to the structures at the intersection points, the structural network of the city is also highlighted with photo-textures to help with the formation of spatial knowledge (Claramunt & Winter, 2007). From the created VEs, two routes of equal size and length that contain the same number of intersection points were selected for the experiment(s) during the encoding stage. Using two similar but not identical routes allowed for repeating the measurements and, thus, increased the internal validity (and statistical power) of the experiment(s). Each route consisted of seven intersection points, with three right and three left turns and one intersection continuing straight. These routes were presented to each participant as a video (thus, passive viewing) of a drive-through to ensure that the duration of the presentation and the content of the visual stimuli were consistent for all participants. These videos were created using ArcGIS[6] software and had the same spatial extent, eye level, and playback speed. The two routes were each presented three times in the three VE types (Abstract VE, Mixed VE, and Realistic VE) in a randomized order. The VEs were projected on a large rear-projection display (2438mm x 1829mm) at a 2.2m distance from the desk the participants were seated at.

3.6 Procedure

The 3D visualization laboratory of the Department of Geography at the University of Zurich was used for the experiments as a controlled environment. After participants arrived at the laboratory, they read and signed the consent form, followed by a brief introduction to the experimental setup and procedure was provided to them. The whole experimental process was performed either in English, or in German, depending on the participant's fluency and preference. At this point, the MMSE was administered verbally (Figure 9). As mentioned in Section 3.5.1, if a participant did not score 27/30, they were to be excluded from the experiment. Since scoring below 27 in the MMSE can be sensitive news to deliver, the exclusion news was prepared with care using a protocol based on advice obtained from the Dynamics of Healthy Aging team at the Department of Psychology (no one scored under 27; thus, this protocol was never used). After discussing any questions that arose, the experiment began: Participants were introduced to the scenario and a passive drive-through of a route was displayed. Immediately then, participants were presented with a series of questions that were designed to require the use of visual, spatial, and visuospatial memory (see Section 3.6.1). While the order of the VEs was rotated using a Latin squares design, the order of the tasks was fixed to counter the learning effect (since the presentation of one task, if presented in a different order, would assist with solving another). Once all the questions for the first VE were answered, participants were provided with the second and third VEs. At this point, a small break was offered, and then the last three videos were presented. Having seen all six videos (two distinct routes for the three VEs), participants proceeded with the last round of visual, spatial, and visuospatial tasks. Once

[6] https://www.arcgis.com/index.html

these were solved, participants were given a 2D paper-map of each of the six VEs from the top-view with the start and end points marked (Figure 12, right). The first part of the experiment ended as soon as participants sketched the route they followed on these 2D maps using pen and paper for all six videos. After one week (6–8 days), participants returned to the lab and repeated the tests without seeing the videos again. Once these were solved, participants performed the standardized psychological tests along with some demographic questions and all the other additional measures (color blindness, expertise levels, sleep, etc. as presented in Section 3.5.1). With this session, the whole experiment was completed. On average, the first experiment (immediate recall stage) took 1h 10min, and the second one after a week (delayed recall stage) took 45 min.

3.6.1 Experimental tasks

The first instruction was intended to enable *intentional* learning: Participants were told to pay attention to learn the route they were about to experience as they would have to re-take the route later on their own. After they experienced the virtual drive-through, they were presented with a series of visual, spatial, and visuospatial tasks. Further details on the sequence and the precise wording of these tasks can be found in the Appendix. Below is an overview of the task definitions (from Publication I, pp.142–143, lightly edited):

- **Visual tasks** can be solved based on photographic information, mainly relying on visual memory: Based on six scenes—three correct (encountered during drive-through), three foils (distractors)—for each VE, participants indicated whether they had *seen the image or not* (Figure 12), using a 6-point Likert scale ranging from "definitely seen" to "definitely not seen". This task was presented in both the immediate and the delayed recall stages.
- **Spatial tasks** can be solved without photographic information: Participants were given their starting orientation (e.g., "you were facing north when you started") and asked to identify the *direction they were facing* at the end of the route, and the *number of turns they took* during the virtual drive. These two questions were asked only in the immediate recall stage. They were omitted from the delayed recall stage as it would be impossible to link the recalled information to the visualization type in the delayed recall stage.
- **Visuospatial tasks** require using a combination of photographic visual details and spatial judgments. This set of tasks consisted of two sub-tasks that differed in the perspective view from which participants solved the task:
    - **'Perspective-preserving' tasks** consisted of the identification of *turn-by-turn recall tasks* on screenshots of the intersections on the route (with a total of seven), all from a first-person perspective. Participants were asked to identify the direction of the turn at these intersection points (Figure 12), as well as to select *the start*- and *end-points* from four options, where only one was correct. These tasks were presented in both recall stages. Identifying the direction of turn is an important signifier for route recall given that participants learned the route from the first-person perspective, i.e., *turn-by-turn recall* tasks were central for the perspective-preserving task type.
    - **'Perspective-switching' tasks** were those where participants needed to mentally switch from the first-person perspective in the encoding phase to an aerial perspective in the decoding phase. This task type also consisted of two subtypes: i) *identification of the route* followed from four aerial-view 2D alternatives, and ii)

map-sketching tasks with active pen-and-paper drawing of the route on aerial-view 2D basemaps. These 2D map views were screenshots of the VEs from the aerial perspective in both sub-tasks and, thus, contained a similar amount of visual realism to the respective VE (Figure 12). *Identification of the route* was only possible in the immediate recall stage, whereas the *map-sketching tasks* were presented in both recall stages, and thus were central to the perspective-switching task type.

Figure 12*: Example scenes presented to participants depicting (i) a visual task as displayed for the Realistic VE, (ii) a turn-by-turn visuospatial task as displayed for the Abstract VE, and (iii) a map-sketching visuospatial task as displayed for the Mixed VE, with the starting and ending points marked on the route.*

Clarification regarding task names in the published papers Note that the naming of tasks may appear slightly different in the included publications. This is because each publication focused on specific task types; thus, the precision of task naming differed. Specifically, in Publication I, the tasks for which the perspective was preserved (turn-by-turn recall and identification of start- and endpoints) were classified and named as visuospatial tasks, while the tasks for which the perspective was changed were named "map tasks" (identification of the correct route and an active reproduction of the sketch). In Publication II, the focus was explicitly on the turn-by-turn recall tasks; thus, when Publication II refers to "visuospatial tasks", it refers to this specific visuospatial task. In Publication III, the focus was explicitly on map-sketching tasks and the wording is similar ("*map-sketching*" or "*sketching*") throughout the paper which, in essence, is a unique instance of a visuospatial task. This manuscript gives an overview of all three publications, thus featuring all of the tasks, therefore, for better distinction, some rewording was necessary.

4. Results

This chapter presents a summary of the key findings included in the publications. In addition, the main effects of each independent variable are provided to give the reader a more comprehensive view of the thesis. The chapter is structured around the following key dimensions (i) *levels of realism*, (ii) *route recall in short- and long-term*, and (iii) *age differences and cognitive abilities.*

4.1 Visualization type: Effects of visual realism on route recall

This section presents the effects of varying visual realism (i.e., *visualization type*) on participants' route recall performance at two levels: First, the results for *all tasks* are presented at an aggregate level for the younger age group to establish a baseline and examine whether the visualization manipulation works as intended. After this benchmarking effort (Publication I), deeper analyses for two of the task types are presented (Publications II and III). For these two *visuospatial* task types, i.e., perspective-preserving *turn-by-turn recall tasks* and perspective-switching *map-sketching tasks*, visualization type had the most consistent effect, driving the rest of the thesis' focus. These tasks are analyzed for both age groups to test the hypotheses related to age. The effects of the *visualization type* for the remaining task types (visual and spatial) are briefly explained, and the reader is referred to the relevant publications. This section (4.1.) answers the first leading RQ:

> **Leading research question (I)** How do varying levels of visual realism in 3D VEs affect the recall of visual, spatial, and visuospatial information in a virtual route learning task?

4.1.1 Recall accuracy

Reporting on participants' overall accuracy for the three *visualization types* over all tasks, Publication I provides a starting point for defining which tasks to investigate in-depth. To obtain a baseline validation on the effects of visualization type, the recall accuracy of the younger group for all *visual, spatial,* and *visuospatial* task types was analyzed. In this analysis, irrespective of task types, participants' recall accuracies differed based on *visualization type:* Participants had the highest overall recall accuracy with the Mixed VE (73.8 %± 11.5) compared to both the Abstract (61.6% ± 12.0) and the Realistic (66.1% ± 11.9) VEs. These differences ($F(2,84) = 21.1$, $p < .001$***, $\eta_p^2 = .154$) had medium to high effect sizes (Mixed-Abstract $p < .001$***, d = 1.03, Mixed-Realistic $p < .001$***, d = 0.65). Moreover, the participants' overall recall accuracy was higher with the Realistic VE than with the Abstract VE with a small to medium effect size (Realistic-Abstract $p < .05$*, d = 0.37). At the task level, these patterns were not consistent (Figure 13). Figure 13 shows that participants' recall accuracy was close to chance level with **visual tasks**, especially with the Abstract VE. The Mixed VE facilitated the highest recall, though this difference was statistically significant only against the Abstract VE with a high effect size. The Realistic VE facilitated a similar recall rate to the Mixed VEs, whereas it was superior to the Abstract VE with a medium effect size. For the **spatial tasks,** participants' recall rates were similar across the three *visualization types* with no statistically significant differences. For the **two visuospatial tasks** (perspective-preserving and perspective-switching), there was a clear statistically significant effect: Mixed VE facilitated higher recall than both the Abstract and the Realistic VEs. Effect sizes were high for the perspective-preserving visuospatial task and small to medium for the perspective-switching visuospatial task. The detailed descriptions of these findings can be found in the Results section of Publication I. These findings guided the selections of tasks for the other two objectives of the thesis, i.e., (i) short- and long-term recall and (ii) group differences. Thus, below the *perspective-preserving* and the *perspective-switching visuospatial recall* tasks are presented in more detail (Table 1). Note that in Table 1, the results were *extended* to include the older adults, as this manuscript aims to examine whether the main effect for the two selected visuospatial tasks remained stable for all participants.

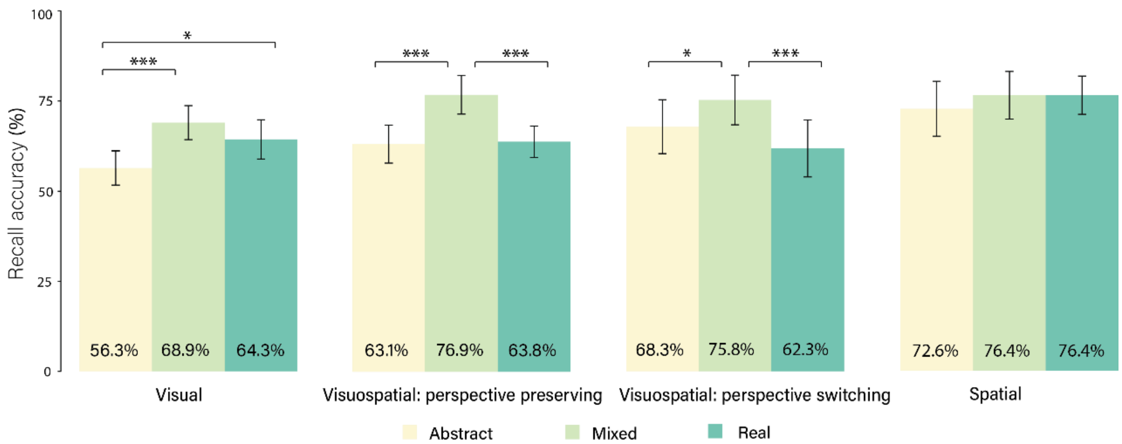


Figure 13*: Participants' recall accuracy rates with the three VEs (left to right) for: Visual, visuospatial: perspective-preserving (turn-by-turn and identification of start- and endpoints), visuospatial: perspective-switching (identification of the route and map-sketching), and spatial tasks. Error bars: ± SEM. ***p < .001, *p < .05. This figure differs from the Figure 4 in Publication I, as it explicitly presents the perspective-switching tasks, and labels are updated to match the wording in this manuscript.*

Table 1*: Participants' mean recall accuracy with the three VEs in (i) the visuospatial turn-by-turn recall task and (ii) the map-sketching task; aggregated for the two age groups in immediate recall stage. ANOVA (F, p, η_p^2) and pairwise comparisons for the statistically significant differences are listed, with the "winning" VE listed first (e.g., M-A means the Mixed VE led to a higher recall than the Abstract VE). SD: Standard Deviation. ***: p < .001, **: p < .01. Results are extended from Publication I (Lokka & Çöltekin, 2019), now including both age groups.*

Task	**Abstract (A)** Mean ± SD (%)	**Mixed (M)** Mean ± SD (%)	**Realistic (R)** Mean ± SD (%)	**Repeated measures ANOVA**	**Pairwise comparisons**
Perspective-preserving: turn-by-turn recall	56.3 ± 22.4	71.3 ± 23.2	61.8 ± 18.6	F(2,160) = 16.5, p < .001***, η_p^2 = .08	**M-A p < .001***, d = 0.65** **M-R p < .01**, d = 0.45** R-A p > .05, d = 0.27
Perspective-switching: map-sketching	55.4 ± 32.2	59.7 ± 31.9	51.1 ± 32.1	F(2,160) = 8.00, p < .01**, η_p^2 = .01	**M-R p < .01**, d = 0.14** M-A p > .05, d = 0.27 A-R p > .05, d = 0.13

As Table 1 shows, the main effects (repeated measures ANOVA column) were statistically significant in participants' recall accuracy across *visualization types*. Pairwise comparisons (last column, Table 1) revealed that the Mixed VE facilitated the highest recall rate with statistically significant differences between the two tasks with moderate to high effect sizes except in one case. For the perspective-preserving (turn-by-turn) task, the Mixed VE facilitated higher recall than both the Abstract and the Realistic VEs, whereas for the perspective-switching (map-sketching) task, Mixed VE was superior only to the Realistic VE. Interestingly, participants' recall accuracy did not differ between the Abstract and the Realistic VE in these two tasks (Table 1).

4.1.2 Visualization type preference

Participants' *visualization type* preference among the three VEs was obtained twice: (i) *before* the experiment, and (ii) *after* the experiment, after having solved all the tasks. Participants marked their preference for one of the three VEs, and their responses were as follows (Table 2):

***Table 2**: Participants' self-reported preference (%) for the three visualizations before and after the experiment.*

Preference	Abstract (A)	Mixed (M)	Realistic (R)
Before experiment	0%	7%	93%
After experiment	4%	54%	42%

Overall, participants strongly prefer the Realistic VE first, but many shift their preference to the Mixed VE ($\chi^2(1) = 55.42$, $p < .001$) after having worked with the VEs. The odds ratio of this shift was at 16.80 (6.879, 47.536), i.e., the odds of participants changing their preference from the Realistic to the Mixed VE were 16.80 times higher than maintaining their Realistic VE preference. None of the participants shifted their preference from the Mixed to the Realistic VE. The preference differences between the two age groups is presented in Section 4.3.

4.1.3 Response confidence: Participants' calibration errors

As an interesting additional dependent variable, participants' ability to calibrate their confidence with their performance, or their "calibration error" (Dodson, Bawa, & Krueger, 2007), was evaluated in Publication II. In this section, the main effects of the *visualization type* on calibration errors are presented. The ways in which calibration errors vary based on the *recall stage* and *age group* are presented in Sections 4.2.3 and 4.3.1. The calibration error, as proposed by Dodson, Bawa, and Krueger (2007), is originally calculated as the division of recall accuracy by participants' confidence ratings. For easier interpretation, the obtained values were scaled to diverge from zero, where values diverging from zero in a positive direction (+) shows overconfidence, and negative (-) underconfidence. The calibration errors of participants for each *visualization type* are shown in Table 3.

***Table 3**: Mean calibration errors for all participants (n = 81) in the perspective-preserving visuospatial task (turn-by-turn recall) in relation to the visualization type. SD: Standard Deviation. Positive values (+) indicate overconfidence, negative values (-) indicate underconfidence.*

Visualization type	**Calibration error** Mean ± SD
Abstract	0.18 ± 0.41
Mixed	-0.04 ± 0.47
Real	0.12 ± 0.36

As Table 3 shows, participants were overall overconfident when using the Abstract and Realistic

VEs. Interestingly, their calibration errors were near zero with the Mixed VE; thus, confidence levels were a near-match with their actual performance (with a slight underconfidence) with the Mixed VE. These observed differences in calibration errors among the three VEs were statistically significant ($F(2,160) = 17.95$, $p < .01$**, $\eta_p^2 = .08$). From the pairwise comparisons, we see that statistically significant differences were driven by the interactions between the Mixed-Abstract ($p < .001$***, $d = 0.51$) and the Mixed-Realistic ($p < .001$***, $d = 0.40$) VEs. The calibration error difference between the Abstract and the Realistic VEs did not yield a statistically significant difference ($p > .05$, $d = 0.15$) where participants were overconfident in both *visualization types*. The calibration errors in relation to the visualization type for the two *age groups* were based on a separate analysis conducted in connection to Publication II, which is summarized in Section 4.3.

4.2 Use context: short- and long-term route recall with the three visualization types

This section presents the effect of lapsing time on how well participants could recall details from the route they learned from the three VEs, addressing the second main variable: *short- and long-term route recall* (i.e., recall stage). This section first provides a summary of participants' recall accuracy for all tasks at an aggregate level in the immediate and delayed *recall stages,* then it elaborates on the two selected visuospatial task types, thus providing answers to the second leading RQ:

> **Leading research question (II)** Do the effects of varying levels of realism in 3D VEs on route recall tasks persist over time?

4.2.1 Effects of recall stage for all tasks

A decline in recall rates over the course of a week was clearly observed at an aggregate level with all *visualization type*s for all participants and all tasks (Table 4, row 1). Table 4 further breaks down this decline in participants' recall accuracy over time for the three *visualization types*. Similar to the previous section, the results shown in Table 4 include all participants (both younger and older age groups aggregated); thus, it extends the findings from Publication I where an overall analysis of the effects of the *recall stage* was performed for the younger participants alone.

Table 4*: All participants' (n = 81) mean recall accuracy for visual, perspective-preserving visuospatial (turn-by-turn recall), and perspective-switching visuospatial (map-sketching) tasks summarized in the two recall stages irrespective of age. Pairwise comparisons for each visualization type (immediate vs. delayed recall stage) are presented in the last column. SD: Standard Deviation. ***: p < .001, **: p < .01.*

Visualization type	**Immediate recall** Mean ± SD (%)	**Delayed recall** Mean ± SD (%)	**Pairwise comparisons** between immediate & delayed recall
All VEs	62.0 ± 14.3	51.6 ± 20.8	**t(80) = 6.39, p < .001***, r = .29**
Abstract VE	58.2 ± 13.6	50.0 ± 19.6	**t(80) = 3.09, p < .01**, r = .25**
Mixed VE	66.7 ± 14.8	56.0 ± 21.9	**t(80) = 3.65, p < .001***, r = .29**
Realistic VE	61.0 ± 13.3	48.8 ± 20.3	**t(80) = 4.50, p < .001***, r = .36**

As demonstrated in Table 4, there were statistically significant differences between participants' recall accuracy in the immediate and delayed recall stages, showing a clear decline visible for all three *visualization types* with strong (Abstract and Realistic VEs) and medium (Mixed VE) effect sizes. *Visualization type* led to statistically significant differences both for the immediate (F(2,160) = 13.4, p < .001***, η_p^2 = .06) and the delayed *recall stage*s F(2,160) = 16.7, p < .001***, η_p^2 = .07). For the immediate *recall stage*, pairwise comparisons revealed statistically significant differences between the Mixed-Abstract (p < .001***, d = 0.60), and the Mixed-Realistic (p < .01**, d = 0.41) with a medium effect size, but not between the Abstract-Realistic (p > .05, d = 0.21) VEs. For the delayed *recall stage*, pairwise comparisons revealed statistically significant differences once again between the Mixed-Abstract (p < .001***, d = 0.29), and Mixed-Realistic (p < .001***, d = 0.34), both with a small effect size, but not between Abstract-Realistic (p > .05, d = 0.06) VEs. *Forgetting rates* are presented in Publication II. In the following pages, similar to Section 4.1, the visuospatial tasks on perspective-preserving (turn-by-turn recall) and perspective-switching (map-sketching) are extensively reported. For the visual and spatial tasks, and the interactions of *age* x *visualization type x recall stage*, the reader is referred to the relevant sections in the attached publications.

4.2.2 Effects of the recall stage: The visuospatial tasks

Participants' recall accuracy differences in the two recall stages for the two visuospatial tasks are presented in Table 5, marking clear differences based on recall stage for each task type.

***Table 5**: Mean recall accuracy of all participants (n = 81) for the two visuospatial tasks: (i) perspective-preserving turn-by-turn, and (ii) perspective switching map-sketching task in the two recall stages, with pairwise comparisons for each visualization type in immediate vs. delayed recall stages presented in the last column. SD: Standard Deviation. ***: p < .001, **: p < .01, *: p < .05.*

	Visualization type	**Immediate recall** Mean ± SD (%)	**Delayed recall** Mean ± SD (%)	**Pairwise comparisons** between immediate & delayed recall
perspective-preserving visuospatial task (turn-by-turn recall)	**All VEs**	63.1 ± 16.6	49.2 ± 16.0	**t(80) = 6.73, p < .001***, r = .60**
	Abstract VE	56.3 ± 22.4	38.3 ± 20.9	**t(80) = 5.72, p < .001***, r = .23**
	Mixed VE	71.3 ± 23.2	63.6 ± 28.2	**t(80) = 2.15, p < .05*, r = .23**
	Realistic VE	61.8 ± 18.6	45.8 ± 19.0	**t(80) = 6.81, p < .001***, r = .61**
perspective-switching (map-sketching)	**All VEs**	55.4 ± 30.0	52.4 ± 33.5	t(80) = 1.71 , p > .05, r = .02
	Abstract VE	55.4 ± 32.2	51.8 ± 33.7	t(80) = 1.54 , p > .05, r = .02
	Mixed VE	59.7 ± 31.9	55.5 ± 35.9	t(80) = 1.73 , p > .05, r = .02
	Realistic VE	51.1 ± 32.1	49.8 ± 34.8	t(80) = 0.53 , p > .05, r = .00

For the **perspective preserving** task, participants' recall accuracy differences were statistically significant in the *immediate recall stage* based on *visualization type* F(2,160) = 16.5, p < .001***, η_p^2 = .08 (see Table 1 in Section 4.1.1). Pairwise, these differences were statistically significant between the Mixed-Abstract (p < .001***, d = 0.65), and Mixed-Realistic (p < .01**, d = 0.45) but not between the Abstract-Realistic VEs. For the *delayed recall stage*, there were also statistically significant differences across visualization types (F(2,160) = 33.3, p < .001***, η_p^2 = .18): Participants remembered more with the Mixed VE compared to both the Abstract (p < .001***, d = 1.02) and the Realistic VEs (p < .001***, d = 0.74) with high effect sizes; and they remembered more with the Realistic VE than with the Abstract (p < .05*, d = 0.38) albeit with a small effect size. For the **perspective switching** task, however, the success rates were generally low in both recall stages (Table 5), and there were no statistically significant differences in participants' recall accuracy between the two *recall stages* irrespective of the *visualization type*. The differences between the three VEs in the immediate *recall stage* (F(2,160) = 8.00, $p < .01^{**}$, $\eta_p^2 = .01$, see Table 1 in Section 4.1.1), were explained by the difference between the Mixed-Realistic VEs ($p < .01^{**}$, d = 0.14), while the others were not statistically significant (Mixed-Abstract $p > .05$, d = 0.27, Abstract-Realistic $p > .05$, d = 0.13). For the delayed *recall stage*, the differences across the three VEs (F(2,160) = 4.7, $p < .01^{**}$, $\eta_p^2 = .00$), were also explained by the difference between the Mixed-Realistic VEs ($p < .01^{**}$, d = 0.16), while others were not statistically significant (Mixed-Abstract: $p > .05$, d = 0.11, Abstract-Realistic: $p > .05$, d = 0.06). More detailed analyses of the effects of recall stage are included in all three publications, including an analysis of forgetting rates.

4.2.3 The effects of the recall stage on participants' response confidence (calibration errors)

The effects of lapsing time (recall stage) on participants' response confidence were also evaluated for each visualization type (Table 6). As Table 6 shows, between the immediate and delayed recall stages, the only statistically significant difference in participants' calibration errors was for the Realistic VE at this aggregate level, where participants felt overconfident in the delayed recall stage.

Table 6*: Mean calibration errors (i.e., response confidence) for the perspective-preserving turn-by-turn visuospatial task for all participants (n = 81) with pairwise comparisons for each visualization type in immediate vs. delayed recall stages. Positive values (+) indicate overconfidence, negative values (-) indicate underconfidence. SD: Standard Deviation. **: p < .01.*

Visualization type	**Immediate recall** Mean ± SD	**Delayed recall** Mean ± SD	**Pairwise comparisons** between immediate & delayed recall
All VEs	+0.05 ± 0.24	+0.12 ± 0.36	t(80) = 1.39 , p > .05, r = .11
Abstract VE	+0.13 ± 0.32	+0.23 ± 0.49	t(80) = 1.42 , p > .05, r = .12
Mixed VE	0.00 ± 0.31	-0.09 ± 0.60	t(80) = 1.19 , p > .05, r = .11
Realistic VE	+0.04 ± 0.32	+0.21 ± 0.38	**t(80) = 3.14 , p < .01**, r = .24**

In the immediate recall stage, there were statistically significant differences in participants' recall accuracy across the three visualization types (F(2,160) = 6.7, $p < .01^{**}$, $\eta_p^2 = .03$). The pairwise comparisons revealed a statistically significant difference between the Mixed-Abstract ($p < .01^{**}$, d = 0.44), the Abstract-Realistic ($p < .05^{*}$, d = 0.31), but not the Mixed-Realistic ($p > .05$, d = 0.13) VEs. Similarly, in the delayed recall stage there were statistically significant differences across the three visualization types (F(2,160) = 15.6, $p < .001^{***}$, $\eta_p^2 = .08$). The pairwise comparisons revealed a statistically significant difference between the Mixed-Abstract ($p < .001^{***}$, d = 0.58) and the Mixed-Realistic ($p < .001^{***}$, d = 0.60), but not between the Abstract-Realistic ($p > .05$, d = 0.04) VEs. The calibration error analysis is further detailed in **Publication II.**

4.3 Effects of age differences and cognitive abilities on route recall

This section focuses on the effect of participant characteristics[7] on the observed dependent variables, specifically *age* and *cognitive abilities*. Due to the thesis' focus on visuospatial tasks, the results included in this section also derive from these tasks. Again, for the visual and spatial tasks, a

[7] The additional measures described in Section 3.5.1 which were collected for control purposes, did not yield any correlations, possibly due to the small sample sizes. Thus, they have not been given further elaboration.

brief overview is provided at the end of the section where the reader is referred to the relevant publications. The results of this section provide answers to the last two leading research questions:

> **Leading research questions (III)**
> 1) How do varying levels of realism in the studied VEs affect the route recall performance of healthy aging older adults and younger adults?
> 2) How do participants' *spatial abilities* and *visuospatial memory capacity* affect route recall with the three VEs?

4.3.1 Effects of age differences on route recall

The effects of aging, or more precisely, the differences between the two tested age groups in the two recall stages, are presented below. Initially the perspective-preserving turn-by-turn task is presented, followed by the perspective-switching map-sketching task.

The differences in recall accuracy between the age groups In this Ph.D. project, *age differences* were examined within the thesis' framework, i.e., for navigational route recall based on differently designed VEs (*visualization type*) and whether the encoded visuospatial information was retained long-term (for a week). Below, a summary of the findings from Publication II is presented, where the focus is on the interactions between *age* x *recall stage* x *visualization type*. As described in detail in Publication II, a 2 (age) x 2 (recall stage) x 3 (visualization type) mixed-design ANOVA revealed statistically significant differences in recall accuracy for all three independent variables. Figure 14 depicts the main effects for: 2a) age; 2b) recall stage; and 2c) visualization type. The interactions between *age x recall stage x visualization type* are presented extensively in **Publication II.**

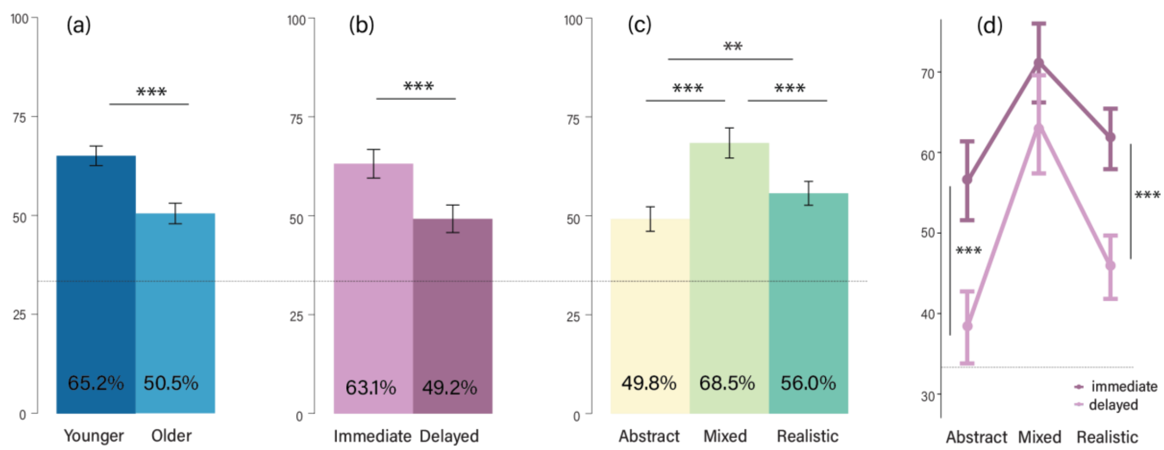


Figure 14*: The main effects of (a) age, (b) recall stage, and (c) visualization type on recall accuracy, and (d) interactions between recall stage x visualization type (irrespective of age). The chance level is marked with a dotted line at 33% for the bar charts. ***p < .001, **p < .01. Error bars: SEM. Reprinted from Publication II.*

Confidence and age: Differences in calibration errors between the age groups The *overall* calibration errors per *visualization type* were presented in Section 4.1. Here, the analysis is extended to the two age groups, as well as the main variable of *recall stage*, in connection with the *visualization types*. As described in Publication II, a 2 (*age*) x 2 (*recall stage*) x 3 (*visualization type*) mixed-design ANOVA revealed significant differences in participants' calibration errors. The main effects are shown in Figure 15 for 4a) *age*; 4b) *recall stage;* (4c) *visualization type*. Also, the interaction between *recall stage* x *visualization type* was statistically significant (Figure 15d).

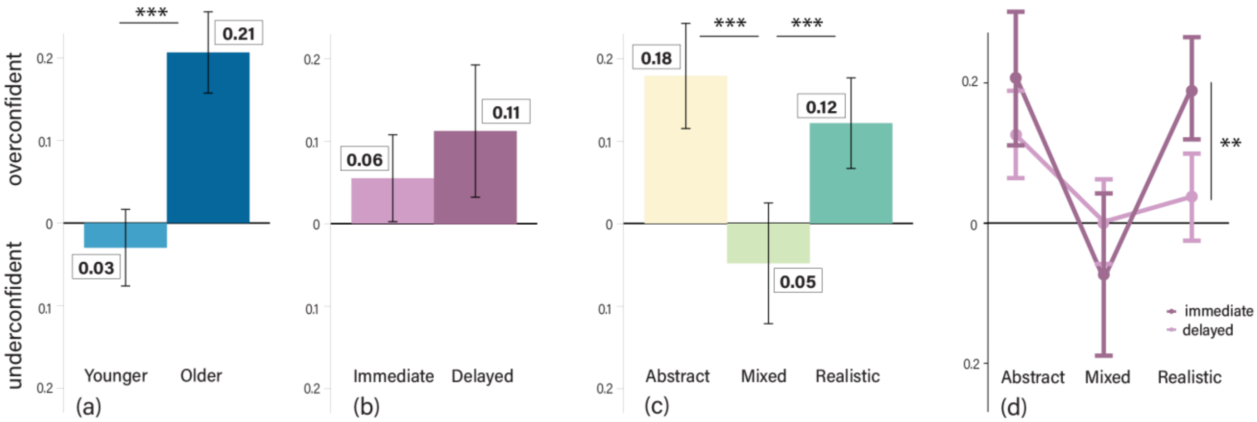


Figure 15*: The main effects of (a) age, (b) recall stage, and (c) visualization type on calibration errors, as well as the (d) interactions between the recall stage x visualization type (irrespective of* age*). *** p < .001, **p < .01. Error bars: SEM. Reprinted from Publication II.*

Similar to the recall accuracy, the interaction between *age* x *recall stage* x *visualization type* is extensively presented in Publication II.

Visualization preference and age: Differences in visualization type preferences After examining the *visualization type* preferences for all participants in Section 4.1.2, this section briefly presents the differences in the preferences of the two age groups (Figure 16).

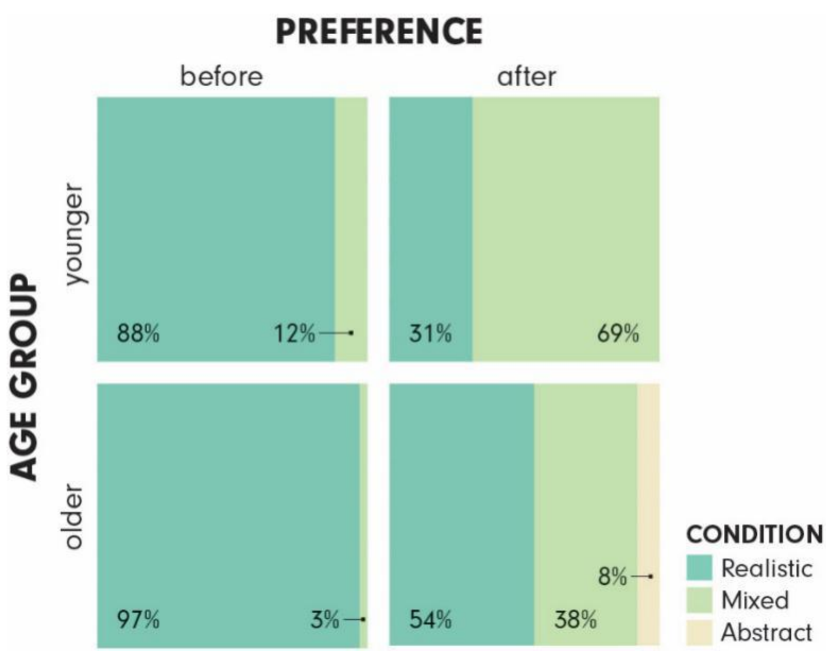


Figure 16*: Visualization type preferences of younger and older participants before and after the experiment. Reprinted from Publication II.*

As Figure 16 shows, *before the experiment*, a clear majority of both the younger and the older participants preferred the Realistic, whereas a minority of the participants preferred the Mixed VE, and *none* of the participants preferred the Abstract VE. *After the experiment*, however, this pattern changes: both age groups shift their preference towards the Mixed VE, with a visible difference between the two age groups in the rate of this change. Detailed statistical analyses and interpretation of these findings can be found in Publication II.

4.3.2 Effects of mental rotation abilities and visuospatial memory capacity on route recall

People's visuospatial memory capacity and spatial abilities differ, and such differences can be important in explaining the variability in their route recall accuracy. To examine if such participant characteristics may explain the variability in addition to age, participants' MRT and VSM scores in connection to the main variables of the thesis were explored. Publication I offers an overview of these differences for younger participants, whereas Publication III examines such differences for the perspective-switching map-sketching task. Below, a summary of the results as featured in Publication III is provided which includes all the main variables for this specific task type.

Interactions with mental rotation abilities (MRT) Participants were split into two groups: Those with higher or lower spatial abilities based on their MRT scores. A 2 (*age*) x 2 (*recall stage*) x 2 (*MRT score*) x 3 (*visualization type*) mixed-design ANOVA revealed that all observed differences in map-sketching performance for all four variables were statistically significant (Figure 17). Statistical analyses and interpretations of these differences are detailed in **Publication III**.

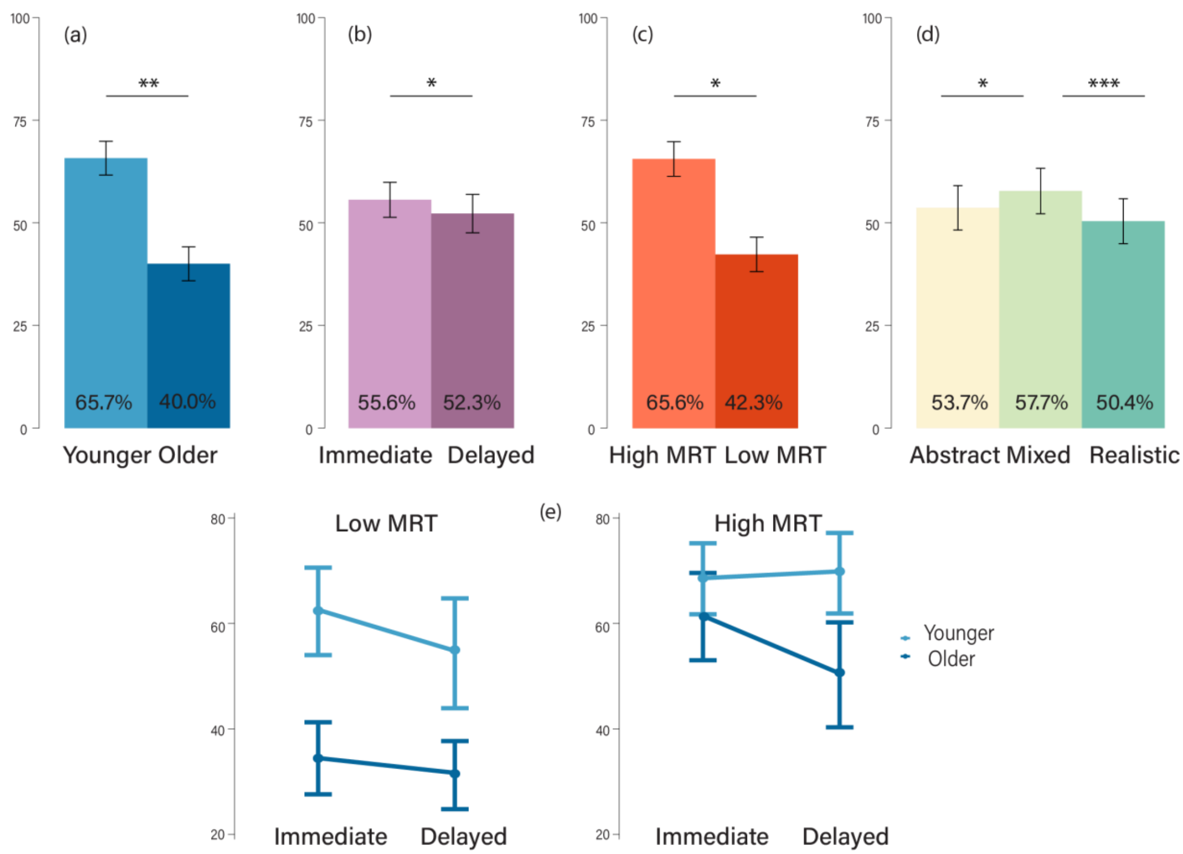


Figure 17*: The main effects of a) age, b) recall stage, c) MRT score, and d) visualization type on the map-sketching task and e) significant interactions of MRT ability x age x recall stage. *** p < .001, ** p < .01, * p < .05. Error bars: SEM. Reprinted from Publication III.*

Interactions with visuospatial memory abilities (VSM) For analyzing participants' recall accuracy differences based on their visuospatial memory abilities, participants were again split into two groups according to their performance on the Visuospatial Memory test ("high" and "low" VSM). A 2 (*age*) x 2 (*recall stage*) x 2 (*VSM score*) x 3 (*visualization type*) mixed-design ANOVA revealed significant differences in the map-sketching performance for three out of the four independent variables: Participants' success in map-sketching did not differ based on *age*, but it did differ based on the *recall stage*, VSM score, and *visualization type* (Figure 18). All these significant differences are extensively reported in **Publication III.** In **Publication I**, Figure 6 demonstrates a summary of the effects of spatial abilities and memory capacity on visual and spatial tasks.

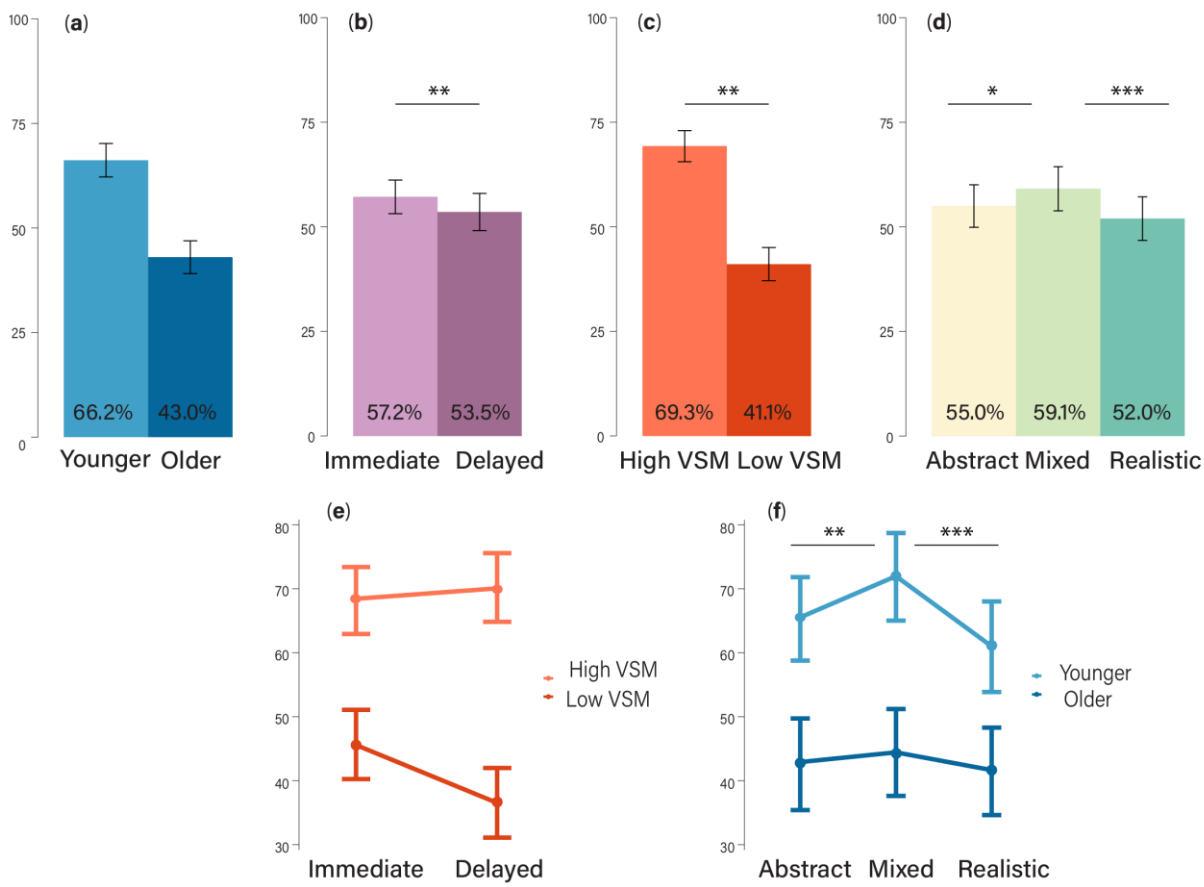


***Figure 18**: The main effects of a) age, b) recall stage, c) VSM scores, and d) visualization type on the sketching task and significant interactions of e) VSM ability x recall stage and f) age x visualization type. *** p < .001, ** p < .01, * p < .05. Error bars: SEM. Reprinted from Publication III.*

5. Discussion of the main findings

This thesis investigated the impact of different levels of visual realism in VEs on the route recall accuracy and confidence of younger and older people, as well as their preferences among the VE options. Three VE designs were varied in their levels of visual realism: one with no photo-textures (the Abstract VE) as a perceptual baseline, one with full photo-textures (the Realistic VE) as the state of the art, and one with photo-textures on selected objects and locations (the Mixed VE) as the "design solution" proposed in this thesis. In a set of controlled laboratory experiments, these three designs were comparatively evaluated in a variety of route recall tasks designed to measure the participants' recall of several types of visual, spatial, and visuospatial information. The results from these experiments allowed for the crafting of answers to the RQs formed under the three dimensions of the thesis' framework. They were based on three main variables: *visualization type*, route *recall stage,* and age-based *group differences* where spatial abilities, and memory capacity were also controlled.

This chapter provides a critical discussion and interpretation of the main outcomes, linking the findings to the original hypotheses and revisiting each RQ, embedding the answers in the relevant literature. At the end of the chapter, a reflection on the limitations of the research is presented.

5.1 The implications of visual realism for virtual route learning

This section (5.1) revisits the thesis' findings linked to the following leading RQ, whereas specific research questions are discussed within the relevant subsections.

Leading RQ (I) How do varying levels of visual realism in 3D VEs affect the recall of visual, spatial, and visuospatial information in a virtual route learning task?

5.1.1 Route recall

Navigational recall is a complex task that requires a combination of cognitive processes, which evolve throughout the lifespan (Klencklen et al., 2012). The overall findings on route recall accuracy in this thesis (irrespective of the *visualization type*) confirm this complexity, but they also demonstrate the importance of visual realism in the visualization design: the route recall performance of an individual is affected by the levels of visual realism as design modifications when they learn from VEs.

RQ1.1. How does Mixed VE affect participants' recall accuracy of visual, spatial, and visuospatial information in route learning tasks when compared to the two baseline conditions (Abstract and Realistic VEs)?

Examining the impact of the **visualization type** irrespective of the other factors (i.e., *age groups*, ability and task variations, and *recall stages*) provides support to the central assumption of this thesis: There is a clear improvement in participants' route recall rates with the Mixed VE. This means that if relevant photographic information is shown in the 'right places' on a route in a VE, it improves route recall. Careful consideration of the two design variables, *i.e.,* the use of realistic photo-textures as a highlighting mechanism (Boér et al., 2013; Borkin et al., 2013; Brady et al., 2008; Isola et al., 2011), and the deliberate selection of the locations of the highlighted landmarks (Claramunt & Winter, 2007; Röser, Hamburger, et al., 2012) are clearly important for successful route learning in VEs. This may be because preserving visual realism *only* in critical locations (Claramunt & Winter, 2007; Klippel et al., 2005; Klippel & Winter, 2005; Lenneberg, 1961; Röser, Hamburger, et al., 2012; Waller & Lippa, 2007) possibly "balances" viewers' cognitive load (Sweller, 1988). This finding is also important as it demonstrates that the infamous "inverted u curve", i.e., the previously documented negative implications of *abundancy or lack* of visual information for human cognition (Smallman & John, 2005a; Yerkes & Dodson, 1908), also applies in a unique new context where different age groups were considered in a virtual route learning task.

Besides the clear overall effects of the Mixed VE, a critical finding in this thesis relates to **the task characteristics**. Overall, participants' performance differences based on task types suggest that higher levels of visual realism facilitate higher recall accuracy than abstract visualizations in tasks that require *visual* information processing, but not in tasks that require predominantly *spatial* information processing. As expected, for *visuospatial* tasks, participants recalled more visuospatial information with the Mixed VE than the other two VEs. For predominantly *visual* tasks, participants could recall more with the Mixed VE than with the Abstract VE, but not with the Realistic VE, suggesting that it is important to assess the value of visual realism for different task types. More specifically, for tasks that rely on photographic information alone, photorealism indeed does *not* impair performance. This finding challenges the naïve realism proposition (Smallman & John, 2005b; Smallman & Cook, 2011) if we take it as a "blanket assumption" against visual realism in displays in general; however, this interpretation should be further confirmed in future studies that specifically control for task-related nuances. For the spatial tasks, there were no differences in participants' route recall performance based on the visualization type. This outcome confirms that for the "predominantly spatial" task participants indeed did not need to make use of photographic information, and demonstrates that task type is critical in assessing a particular visualization type or design. Thus, the findings partially confirm hypothesis **H-RQ1.1.A:**

> **H-RQ1.1.A. The Mixed VE will be superior to the other two VEs in facilitating participants' route recall accuracy across visual and visuospatial recall tasks that require the recall of some visual cues, whereas for spatial tasks, photorealism should not offer an advantage**: Thus, irrespective of their age, participants' recall performance across visual and visuospatial recall tasks will be best with the Mixed VE.

Route learning theories suggest that people learn a route by associating directional information with the landmarks they pay attention to (Waller & Lippa, 2007; Zhong & Moffat, 2016). This thesis shows that taking important properties that define a landmark (persistence, perceptual saliency, and informativeness) (Stankiewicz & Kalia, 2007) and transforming those into a visualization type (Mixed VE) can have profound effects on route learning performance and a person's recall of the learned route. When these landmarks are placed in the "effective landmark locations" that are critical for navigation (Röser, Hamburger, et al., 2012), the resulting highlighted VE unsurprisingly facilitates better learning in tasks that require the recall of photographic cues. However, as mentioned above, this is clearly expressed in the results linked to the visuospatial tasks, whereas it is not evident in visual or spatial tasks alone. While it is possible that "visual realism is not too bad for visual tasks", as argued above, this may also be due to the fact that, as opposed to the Mixed VE, in the Realistic VE the photo-textures were not controlled. Photographic information appeared as it would normally appear in a real-world city. Therefore, participants could have paid attention to any of the photo-textured structures that had nameable elements and colors more memorable to them (Brewer, 1996; Brown et al., 2011; Lenneberg, 1961; Özgen, 2004). The reason for the lack of difference in participants' performance in the spatial task across visualization types might be that, for such tasks, photographic information may be simply irrelevant, therefore the Abstract VE is able to support identifying one's cardinal orientation or the number of turns one takes.

Perspective-preserving turn-by-turn visuospatial task: In this task, the viewing perspective was held constant during encoding and decoding, and a consistent pattern was observed across the age groups' route recall performance in favor of the Mixed VE, confirming the hypothesis **H-RQ1.1.B**.:

> **H-RQ1.1.B. The Mixed VE will be superior to the other two VEs in facilitating participants' route recall accuracy in perspective-preserving visuospatial tasks:** As the critical visuospatial information for route recall (Röser, Hamburger, et al., 2012) remains present and highlighted in the Mixed VE during both the encoding and decoding of information, it will help participants, irrespective of their age, to identify the direction of the turn at intersections better than the other VEs.

Perspective-switching map-sketching visuospatial task: For the tasks that required switching from a first-person view in the encoding phase to an aerial view in the decoding phase, there was only partial evidence to support the proposition that the Mixed VE assists spatial knowledge acquisition and retention better than the others. Participants' aggregate recall accuracy with the Mixed VE was better *only* with the Realistic VE. Descriptive statistics suggested that the Mixed VE facilitated a higher recall accuracy than the Abstract VE, but this difference was not statistically significant. Thus, these results only partially confirm the hypothesis **H-RQ1.1.C**:

> **H-RQ1.1.C. The Mixed VE will be superior to the other two VEs in facilitating participants' route recall accuracy in perspective-switching visuospatial tasks**: The highlighted elements in the 2D basemap of the Mixed VE will serve as anchoring points to assist people, irrespective of age, to perform the perspective switch (Thorndyke & Hayes-Roth, 1982).

The outcome that the participants' recall accuracy with the Mixed VE was not higher than with the Abstract VE for this task was surprising. In this task, participants' aggregate recall accuracy was approximately 60%, confirming the known difficulty of tasks that require perspective-switching (Golledge et al., 1995; Taylor & Tversky, 1996). This difficulty, up to a point, may have occluded some statistical interactions (i.e., cause a Type II error) at the aggregate level. However, evaluating participants' actively produced sketches based on accuracy and completeness showed that their success rates were above the chance level (Montello, 1998). Perhaps the cognitive load (Sweller, 1988) induced by this difficult task that required a mental rotation was already high, and the additional cognitive load induced by the photographic information in the Mixed VE further impaired participant success. It is important to remember that people may not necessarily need visual cues to build a 2D map of a followed route if they imagine a 2D route of the path in their minds as they take turns.

RQ1.2. How do the Abstract and Realistic VEs differently affect participants' recall accuracy of visual, spatial, and visuospatial information in route learning tasks compared to each other (Abstract VE vs. Realistic VE)?

For RQ1.2 results point to a mixed pattern: At the aggregate level for all tasks, participants' recall rates with the Realistic VE were higher than the Abstract VE. For the visual tasks, this outcome was the same. For the visuospatial and spatial tasks, however, this pattern was no longer present, *i.e.,* there were no differences between the two VEs. In other words, as soon as there was a *spatial* component in the task, the Realistic VE was no different than the Abstract VE. In sum, in this study the evidence is mixed in testing the hypothesis **H-RQ1.2.** for the superiority of the Realistic VE to the Abstract VE; meaning that the hypothesis can be retained at the "overall recall performance" but at the task level this claim holds only for visual tasks:

> **H-RQ1.2. Participants' *overall recall performance* with the Realistic VE will be better than with the Abstract VE:** As the Realistic VE provides more visual cues that include human recognizable and nameable elements known to support recall (Borkin et al., 2013; Brady et al., 2008; Isola et al., 2014).

This finding is important because it contributes new knowledge to the theories investigating abstraction and realism, and it further confirms the arguments presented for H-RQ1.1.A regarding task type. The nuanced outcomes regarding task type highlight the importance of reflecting on the characteristics of the recall task and clearly defining it in similar experiments. In this thesis, for a *visual* recall task, the added visual cues indeed support recall, improving individual accuracy as seen from other research (Brady et al., 2008; Potter & Levy, 1969). In tasks that are fully or partially spatial, such as visuospatial and spatial tasks, Realistic VE does not offer any advantages to the Abstract VE, which suggests that the design needs are different for different tasks. In the two perspective-preserving visuospatial tasks (turn-by-turn and map-sketching), there was a demand for recalling not only visual *but also* spatial information, thus possibly increasing the cognitive load. In the perspective-switching map-sketching task, the required mental rotation (as mentioned earlier) may have led to additional task complexity (Golledge et al., 1995). The extra cognitive demand coming either from the lack or the abundancy of visual information (respectively, as in the Abstract and the Realistic VEs) can have detrimental effects on the visuospatial information recall. With the Abstract VE, people may have needed to pay extra attention to identify anchoring points in an environment with scarce visual information, whereas with the Realistic VE they may have needed to extract and isolate the relevant information from an abundantly rich visual representation.

Summary: All findings in response to **RQ 1.1.** suggest that the Mixed VE has a clear positive effect on route recall performance, evidenced by consistent patterns across the evaluated tasks. This confirms the importance of design decisions for efficient route learning with virtual displays.

5.1.2 Preference

In addition to performance measures, user preferences are important in visualization research. They might indicate a cognitive bias and affect people's decisions, even if a visualization assists users more effectively. If users do not like it and are unwilling to use it, they might select a "bad tool" instead. To examine user preferences in this thesis, the following RQ was developed:

RQ1.3. How do different levels of realism affect participants' visualization type preference among the three VEs for route learning *before* and *after* the learning tasks are performed?

Participants' overwhelming preference for the Realistic VE before the experiment, which was subsequently compounded by the clear aversion towards the Abstract VE, confirms the previously documented desirability of visual realism in displays (Smallman & John, 2005b; Smallman & Cook, 2011) in a new context (navigation) and a new display type (virtual reality). The hypothesis **H-RQ1.3.**, therefore, is confirmed:

H-RQ1.3. Participants will prefer the Realistic VE before the experiment: All (both older and younger) participants will prefer the Realistic VE to the other two VEs before the experiment (Smallman & John, 2005b; Smallman & Cook, 2011).

The VE preference *after* participants solved the tasks favor the Mixed VE. However, this preference shift comes with nuance. It is discussed through the lens of group differences in Section 5.3.1.

5.1.3 Confidence

Another important subjective measure in learning tasks is a person's degree of confidence in how well they have learned, especially if they may take action afterwards, such as in route learning. In this thesis, participants' "calibration errors" were used as a confidence measure documenting the differences in achieved vs. perceived accuracy. The RQ on this subject was as follows:

RQ1.4. How does the Mixed VE differently affect the alignment of participants' response confidence with their recall accuracy when compared to the two baseline conditions (Abstract and Realistic VEs)?

Overall, participants are much more aware of how well they have memorized a route they learned with the Mixed VE than with the others. Equally important, they know when they do *not* remember the route, or a given location on that route. With the Mixed VE, participants achieved- and perceived-accuracy are an almost perfect match, thus the original hypothesis **H-RQ1.4.** is retained:

H-RQ1.4. The Mixed VE will be superior to the other two VEs in facilitating the *alignment of confidence with recall accuracy.* Balanced information and reduced visual clutter (Rosenholtz et al., 2007) will assist participants in identifying places they can and cannot recall, better aligning their confidence with their performance.

The healthy confidence levels with the Mixed VE suggest that realistic visual cues are critical for a person to identify that they *can* or *cannot* recall a scene or a direction they took at an intersection point. On the other hand, with both the Abstract and Realistic VEs, participants were *overconfident.* Arguably the Abstract and Realistic VEs contain a certain degree of danger for route learning: If users prematurely believe that they have "learned" the route, they may take actions that put them at risk. An overconfident user may experience unpleasant situations such as getting lost or being late, which can be particularly stressful for older adults. This new finding shows how visualization design can affect individual confidence. Thus, it has potentially important implications for well-being during real-world navigation.

Summary: The better alignment of achieved and perceived recall accuracy with the use of the Mixed VE is a promising finding that is crucial for the development of VEs used for route learning purposes (Richardson, Montello, & Hegarty, 1999). Being able to evaluate one's own learning accurately is a personal skill that is valuable also in other contexts, and if the design of the visualization enhances metacognition, this could lead to better strategies for learning in any context where visualizations are used.

5.2 The implications of visual realism for route learning over time

This section (5.2.) provides answers to the following leading RQ and related specific RQs:

Leading research question (II) Do the effects of varying levels of realism in 3D VEs on route recall tasks persist over time?

5.2.1 Route recall

The first specific RQ investigated the overall effects of visual realism with the lapse of time on participants' route recall accuracy, *i.e.,* immediate and delayed *recall stages*:

RQ2.1. How do different levels of visual realism affect participants' recall accuracy of visuospatial route information in the *immediate* and *delayed recall stages*?

When examining the effects of passing time on participants' route recall accuracy, not surprisingly, a decline in performance was evident: Irrespective of the *visualization type*, participants' recall success decreased from the immediate to the delayed recall stage (a week later). These results confirm the well-established theories suggesting that only a subset of the information initially stored in short-term memory transfers to long-term memory (Byrne, 2017). What remains of the information is permanently lost. When visualization types are examined comparatively, again, Mixed VE is superior to the other two: Participants' recall accuracy with the Mixed VE is higher than both with the Abstract and the Realistic VEs in *both* recall stages. These findings confirm the hypothesis **H-RQ2.1.**:

H-RQ2.1. All participants' recall performance will be best with the Mixed VE, also in the delayed recall stage: Due to the fact that these human recognizable elements assist recall also in the long-term (Brady et al., 2008).

The fact that people could better retain visuospatial information in their memories long-term with the use of the Mixed VE when compared to the other two VEs provides additional support for its potential use as a training device in the route learning context. Route learning, (extensively discussed in RQ2.2.) is heavily dependent on retention over time.

The second specific RQ is an extension of RQ2.1., and examines the *amount of information* that is transferred from short- to long-term memory with the use of the three *visualization types*:

RQ2.2. How do different levels of visual realism affect the amount of visuospatial information transferred from short- to long-term memory, i.e., what are the 'forgetting rates' in visuospatial route recall tasks from the immediate to the delayed recall stages?

In answering this question, an interesting and exciting finding was at the task level: For the **perspective-preserving turn-by-turn visuospatial task**, participants' recall accuracy declined from the immediate to the delayed *recall stage* for the Abstract and the Realistic VEs but *not* for the Mixed VE. Therefore, the hypothesis **H-RQ2.2.** is retained:

H-RQ2.2. The differences from the immediate to the delayed recall stage will be smaller with the Mixed VE: The combination of the balanced cognitive load (Sweller, 1988) and the presence of nameable and memorable realistic elements (Borkin et al., 2013; Isola et al., 2014; Brady et al., 2008) will support transferring higher amounts of visuospatial information to the long-term recall.

This finding is very encouraging for the development of visualizations used for route learning. Essentially, this finding provides evidence that a visualization design created for route learning *can* overcome the known barrier of human cognition in transferring information from short- to long-term recall. Overcoming this known decline in recall over time (Byrne, 2017) might eventually lead to more stress-free navigation experiences in unfamiliar environments, which is specifically important for older adults. An additional finding regarding the recall stage was that participants remembered their route better a week later if they learned it with the Realistic VE than with the Abstract VE. Participants' performance with the Abstract VE was only slightly higher than the chance level. This result highlights a floor effect for the Abstract VE over time, a fact that is not evident for the other two VEs (Mixed and Realistic) and is comparatively demonstrated for the first time in the scope of this thesis. While experiencing a drive-through in VEs that included partial or full visual realism, the fact that (some) visual elements are presented with high realism may assist people in creating mental episodes of the event, potentially associating them with their own prior experiences. Such experiences can then be linked to time and place, helping with the formation of an episodic memory (Gras et al., 2013; Tulving, 1972). The triggering of such episodic events using visual cues may make it easier for people to re-trace their memories even after the lapse of time. While realism appears to be positive based on this interpretation, it is important to note that the Mixed VE facilitated route recall better than the Realistic VE, possibly due to higher cognitive load with the Realistic VE also in the long-term recall. In the Mixed VE, only a selection of structures was realistically presented, and because these were highlighted and emphasized, it may have been easier for people to process this smaller amount of visual information (Cowan, 2010; Miller, 1956) located at important positions (Röser, Hamburger, et al., 2012) and, to later decode it.

For the **perspective-switching map-sketching visuospatial task,** at the aggregate level, the findings differed from the turn-by-turn visuospatial task discussed above (see Table 5): There was *no effect* of elapsed time *(recall stage)* on participants' route recall, irrespective of *visualization type,* suggesting that there must be additional factors interfering with recall over time with this task. A potential explanation might be related to the nature of the task. Map-sketching required an active effort from the participant beyond just recalling the information. They were required to actively reproduce a sketch of the route on a 2D basemap (Mastin, 2010). While the overall performance was not high (55.4% in the immediate, 52.4% in the delayed recall stage, see Table 5), using a pen and actively drawing the route on paper in the immediate recall stage may have supported retention for those who were able to do the task well the first time they tried. In other words, this active effort may have ensured that the information was stored in long-term memory (Dubuc, 2002; Modigliani & Seamon, 1974). While there may be also other factors involved here, this interpretation suggests that a "recipe" to efficient route learning may at least partially lie in ensuring participants' active involvement in supporting cognitive mechanisms for long-term retention. A more fine-grained analysis of this task with the interacting variables based on Publication III is discussed later in this chapter (Section 5.3.2).

For the **visual task,** a severe decline in participants' recall accuracy for the Realistic VE was observed, and slightly less for the Mixed VE, decreasing to the chance level in the delayed recall stage. With the Abstract VE, there was no decline. However, participants' recall accuracy was already close to the chance level in the immediate recall stage. Overall, the Mixed VE facilitated the highest recall accuracy in both recall stages, signifying that selective highlighting with visual realism compensates, up to a point, for the severe difficulty in identifying scenes one experienced a week ago. Even though the findings suggest only a slight advantage when working with the Mixed VE for visual tasks for long-term retention, findings still encourage a hypothesis that Mixed VE (or a similar solution with some more adjustments), might support long-term retention better than a fully Realistic VE. A future study specifically designed to assess this hypothesis could test this position.

5.2.2 Confidence

Participants' confidence was evaluated also in connection with the two *recall stage*s besides the visualization types (reported in Section 5.1.3). The following RQ was developed:

> **RQ2.3.** How does the Mixed VE differently affect the alignment of participants' response confidence with their recall accuracy when compared to the two baseline conditions (Abstract and Realistic VEs) in the *immediate* and *delayed* recall stages?

The results from both the immediate and the delayed recall stages further confirm the Mixed VE as a "safer" visualization type, compared to the Abstract and the Realistic VEs: Participants achieved a perfect match between their perceived and achieved recall accuracy during the immediate recall stage with the Mixed VE, while they were overconfident in their responses with the other two VEs. For the delayed recall stage, the participants were slightly underconfident in their responses with the Mixed VE, while for the Abstract and Realistic VEs, they were even more overconfident. The superiority of the Mixed VE confirms the hypothesis **H-RQ2.3.** stating:

> **H-RQ2.3. The Mixed VE will be superior to the other two VEs in facilitating the *alignment of confidence with recall accuracy* in both recall stages:** Balanced information and reduced visual clutter (Rosenholtz et al., 2007) together with the fact that human recognizable elements can persist in long-term recall (Brady et al., 2008) will assist people in identifying the places they can and cannot correctly recall, thus better aligning their confidence with their performance *also* in the delayed recall stage.

Being able to correctly assess one's own recall performance after a lapse in time further confirms that participants did not only *memorize*, but indeed possibly *learned* the route, and were thus able to correctly assess their own performance a week later. This metacognitive success is clearly evident with the Mixed VE and not with the others, which supports the position that the proposed design intervention works well, from multiple perspectives.

5.3 Implications of visual realism based on age differences and cognitive abilities for route recall

In this section, implications of visual realism for route recall accuracy is discussed from the "group differences" perspective, people with different cognitive abilities due to age, spatial abilities, and memory capacity. This section (5.3) provides answers to the two leading RQs below, whereas the specific RQs are discussed in following sub-sections.

Leading research questions (III)

1) How do varying levels of realism in the studied VEs affect the route recall performance of healthy aging older adults and younger adults?
2) How do participants' *spatial abilities* and *visuospatial memory capacity* affect route recall with the three VEs?

5.3.1 Route recall and age differences

The negative effects of aging on cognitive abilities are well-documented in psychology and the related domains (Anders et al., 1972; Foos, 1989; Park et al., 2002). It is also well-understood that age-related cognitive decline negatively affects navigation performance (Meneghetti et al., 2014; Moffat et al., 2001; Muffato et al., 2015). The results from this thesis confirm that, overall, older participants' route recall performance is poorer than that of the younger. This finding in turn confirms the decline in older people's associative learning (Head & Isom, 2010; O'Malley et al., 2018). Specific to this thesis, findings also show how visualization types varying in the amount of visual realism differently affect this decline in cognitive performance caused by natural aging. This new understanding regarding visualization design can lead to professionals providing better support to older adults through custom visualizations. In this context, the following RQ was posed:

RQ3.1. How does the route recall accuracy of older adults differ from that of younger adults with the three VEs?

As discussed in Section 5.1, the Mixed VE facilitated the highest recall when the two age groups were analyzed together. A more in-depth analysis showed that *age x visualization type* x *recall stage* did *not* interact. The lack of interaction among these three variables could be interpreted as a positive outcome for the Mixed VE. In other words, the Mixed VE facilitated better recall than the other VEs *irrespective* of *age groups* or *recall stage* (Figure 3, Publication II). Therefore, the Mixed VE appears to be the "best" among the tested VEs for participants in both *age groups* for the perspective-preserving turn-by-turn recall of visuospatial information. Therefore, the hypothesis H-RQ1.1.B. is also confirmed for age differences. Since the Mixed VE helps both age groups, one can argue that it acts as a cognitive amplifier for the younger, and as a cognitive prosthetic for the older (Arias-Hernandez et al., 2012; Card et al., 1999).

As mentioned earlier, the literature suggests that the perspective-switching visuospatial tasks are particularly difficult for older people (Fricke & Bock, 2018). Findings from this thesis further confirm that the tasks that require perspective switching are indeed more difficult for older people. This is even more evident when the participants are grouped based on their spatial abilities and their visuospatial memory capacity. Thus, hypothesis **H-RQ3.1.** is confirmed:

> **H-RQ3.1. Age and recall**: Younger participants' route recall accuracy will be higher than older participants' irrespective of the VE type (Fricke & Bock, 2018; Moffat et al., 2001; Muffato et al., 2015).

A detailed discussion of the effect of age on participants' sketching performance is provided in Section 5.3.2, where spatial abilities and memory capacity are also included.

The second age-specific research question was on participants' confidence:

> **RQ3.2.** How do different VEs affect younger and older participants differently in self-assessing how well they performed in the given tasks?

One of the most-striking positive effects of the Mixed VE appeared in the examination of "calibration errors" (i.e., differences in response confidence) of the different *age groups*. The calibration error patterns change when the two age groups are separately studied: Table 3 shows the aggregate analysis, whereas Figure 15 shows them separately. Overall, the younger group was accurate in assessing their own performance with a slight underconfidence, whereas the older group was clearly overconfident. These results are in line with the previous work in literature. It has been shown that older people believe they perform better than they actually do, especially in tasks that involve memory (Dodson, Bawa, & Krueger, 2007). However, the Mixed VE supported their metacognition: The older participants' perceived task accuracy was much closer to their real task accuracy, not only in the immediate *recall stage* but also in the delayed recall stage. The younger participants exhibit a curious pattern: They are slightly underconfident in the immediate recall stage, and rather strongly underconfident in the delayed recall stage with the Mixed VE. Based on these findings the hypothesis **H-RQ3.2.** is confirmed:

> **H-RQ3.2. Age and confidence**: Overall, older participants will be overconfident in their responses in comparison to younger participants (Dodson, Bawa, & Krueger, 2007). The Mixed VE should moderate this effect (similar to H-RQ1.4) and the effect should be more pronounced for the older group.

Once again, these findings are encouraging for the potential future use of the Mixed VE as a training device. This *visualization type* seems to enable people to overcome another known difficulty that older people experience, i.e., accurately rating their performance in memory tasks (Dodson, Bawa, & Krueger, 2007). The Mixed VE can provide the necessary assistance to better self-assess one's own performance, and, for example, encourage individuals to seek more training trials when needed. The younger group's confidence levels seem exactly as expected in the immediate recall stage. However, interestingly they are underconfident with the Mixed VE in the delayed recall stage. Underconfidence can be viewed "harmless" compared to overconfidence: If people are underconfident, they could repeat the training until they feel comfortable with it. Taken together, as mentioned earlier, it is plausible to claim that these findings provide further evidence that the use of the Mixed VE can act as a cognitive amplifier for younger people while simultaneously serving as a cognitive prosthetic for older people (Arias-Hernandez et al., 2012; Card et al., 1999).

The third specific RQ investigated the effect of age on participants' preferences:

RQ3.3. How do participants' visualization type preference differ based on age for the three VEs for route learning tasks *before* and *after* the VE experience?

As discussed in Section 5.1.2, the vast majority of all participants preferred the Realistic VE prior to experiencing the virtual drive-through. Participants' preferences shift *after* the experiment suggesting that a large proportion of them realize the value of the Mixed VE and adjust their preferences after being exposed to all VEs. This shift in preference is more pronounced for younger people. This finding is in alignment with what was proposed in hypothesis **H-RQ3.3.**, and thus this hypothesis is also confirmed:

H-RQ3.3. Age and VE preference: *Before* the experiment, all participants will prefer the Realistic VE. *After* the experiment, more of the younger participants (compared to the older) will change their preference for the Mixed VE. Older participants, however, due to an overall decline in their spatial abilities (Moffat & Resnick, 2002), might not be able to identify which visualization supports them better and, thus, will still prefer the Realistic VE after the experiment (Smallman & John, 2005b; Smallman & Cook, 2011).

The overall finding partly confirms and partly contradicts the naïve realism idea (Smallman & John, 2005b; Smallman & Cook, 2011). The fact that the majority of the older participants still preferred the Realistic VE after the experiment suggests a level of naïve realism in older people not found in the younger participants, most of whom switched their preference to the Mixed VE after the experiment. In other words, older people are less able to identify their learning "gain" with the use of the Mixed VE. This metacognition issue further confirms the observations in the confidence analysis presented in this thesis (Dodson, Bawa, & Krueger, 2007). Smallman and colleagues have shown that people with lower spatial abilities have difficulty in adjusting their preference when comparing realistic and abstract display types (Smallman & John, 2005b; Smallman & Cook, 2011).

Summary: The majority of older people may still prefer and thus select and use, a Realistic VE without realizing the harmful effect it may have on their performance. These issues in confidence and preference are important for designers to be aware of and consider when designing VEs.

5.3.2 Route recall and spatial and memory abilities

Navigational performance is shown to be related to individual differences (Ishikawa & Montello, 2006). Furthermore, Smallman and colleagues have shown that naïve realism is stronger in people with lower spatial abilities, as they have difficulty adjusting their preference after working with different display types (Smallman & John, 2005b; Smallman & Cook, 2011). To better understand the interactions between cognitive abilities and age, group differences based on spatial abilities and memory capacity have also been investigated in this thesis. The specific RQ regarding these participant characteristics was formulated as follows:

RQ3.4. How do participants' spatial abilities as measured by mental rotation task (MRT), and visuospatial memory capacity as measured by the visuospatial memory capacity test (VSM) interact with their route recall accuracy with the use of the three VEs taking also age into account?

The thesis' findings demonstrate that spatial abilities and the visuospatial memory capacity of the participants' affect their route recall accuracy. Participants were split based on their MRT or VSM scores, and in both cases the split revealed differences in their overall recall accuracy irrespective of visualization type and age. These findings confirm the hypothesis **H-RQ3.4.A.** stating that:

H-RQ3.4.A. Spatial abilities and memory capacity: Participants with higher Mental Rotation Test (MRT) or Visuospatial Memory Test (VSM) scores will outperform the participants with lower MRT/VSM scores, irrespective of age or VE type (Muffato et al., 2017; Wolbers & Hegarty, 2010).

The findings first of all confirm the usefulness of these tests for predicting spatial abilities and visuospatial memory capacity in a map-sketching visuospatial task that requires (i) a mental rotation to translate information acquired from a first-person view to an aerial 2D map view, and (ii) the retention of visuospatial cues in memory to accurately identify the structures on the basemap for sketching. All participants, irrespective of their spatial abilities and visuospatial memory capacity, performed similarly with the three *visualization types*: For all of them, the Mixed VE facilitated the highest route recall rates compared to both the Abstract and the Realistic VEs. These findings highlight the consistency across people with different sets of abilities marking the fact that a design solution such as the Mixed VE does not only support one type of user (high- or low- abilities), but all of them collectively. For the VSM group, there were no differences for *VSM score x visualization type*, which was somewhat surprising. A reasonable expectation would be that the people in the higher VSM group better cope with the additional information provided in the Mixed and Realistic VEs, as suggested in the first part of the hypothesis below. Thus, the first part of the hypothesis **H-RQ3.4.B.** is rejected. What can be suggested, however, is that the VSM test is a good test for predicting map-sketching performance irrespective of the *visualization type*. On the other hand, there were no interactions across *MRT score x recall stage* x *visualization type*, which indirectly indicates that people with high spatial abilities *do* perform better than people with lower spatial abilities with the use of Abstract VEs. This is possibly linked to the idea that high-spatial people may be better at performing mental rotations without the existence of visual cues to assist them. This leads to the confirmation of the second part of the hypothesis **H-RQ3.4.B.**:

H-RQ3.4.B. Memory capacity and visual cues: Irrespective of age, participants with a higher VSM will outperform the participants with a lower VSM in producing accurate sketches, particularly with the Mixed and Realistic VE, as these provide potentially helpful photorealistic cues (see H-RQ1.1.C) (Vandenberg & Kuse, 1978), whereas MRT will be most relevant to the Abstract VE because this visualization type contains no photorealistic cues (Ekstrom et al., 1976).

In that case, however, people with lower spatial abilities should also have performed better with the use of the Mixed and/or Realistic VEs, given that the added visual cues could act as assistants for them, a fact that was not observed.

Summary: When the two *age groups* were analyzed, there was a difference between the two *recall stages* based on their MRT scores: The higher-MRT younger group did very well after a week, in contrast to the higher-MRT older group. This finding suggests a limit in the amount of information older people can recall after a lapse in time even with higher MRT abilities, a fact that does not exist for the younger group.

5.4 Limitations

As in every scientific study, the contributions of this thesis are accompanied by a set of limitations regarding its scope, methods, and findings. These limitations are acknowledged and reflected upon critically in this section. The limitations are grouped according to the three dimensions as presented earlier: (i) *visualization type*, (ii) *use context*, and (iii) *user characteristics*.

Visualization type/design

- ❖ The Mixed VE combined two parameters to achieve its success: (i) the use of realism as a highlighting mechanism, and (ii) the intentional selection of the locations of the highlighted features based on navigational theories (at the intersections and direction of turn). These two parameters were treated in combination. One might argue this as a possible "confound", because from our findings, it is impossible to determine how these two factors might have affected the route recall performance differently if they were treated independently from each other. In fairness, one simply *has to* apply the photo-textures *somewhere*, thus this is not a completely independent issue. Nonetheless, at this point it is not clear if the photo-textures were to be applied in "random" locations, how much of a difference they would make.
- ❖ One can also argue that the contents of the photo-textures should matter, and indeed they do. This was counterbalanced in this project to some degree, based on qualitative reasoning and relevant literature, for the "nameability" of objects or colors, and whether they had labels or distinct shapes (see Lokka & Çöltekin, 2017). More work is needed to better assess the effects of the texture content.
- ❖ Linking to the above point, it is not clear if a highlighting method other than photo-textures would yield the same success rates. Photo-textures were selected due to their resemblance to the real world and might create navigational anchoring points from VR to the real world since they represent state-of-the-art VR practices. However, manipulating contrast or outlining the features may also improve route memorability.

Use context

- Current results are based on a laboratory study. At this point in time, it is not clear what the implications of these findings are in the real world. More research is needed to better understand the relationship between learning from a VE as a training device and how this corresponds to the real-world performance.
- The represented environment used throughout the project was heavily urban, with building textures similar in shape and size, with nothing *really* standing out, e.g., there were no global landmarks (Credé et al., 2020). While this uniform design was intentional in the presented work to minimize the possible confounding effects of global landmarks, this may have constrained the participants' overall recall performance, as they were not provided with any extra assistance that could have proven helpful in other VE or real world settings.
- The fact that this thesis focused on visuals alone for route learning (no sounds, no tactile feedback from movement, no interaction, only passive watching from the passenger seat) could have limited the amount of information people could recall. As Gale, Golledge, Pellegrino, and Doherty, (1990) report, the value of proprioceptive and/or efferent information is not to be neglected. People who experience a route via walking can better re-travel the route compared to those experiencing it as a video. More specifically, visual information from VR may not be enough to induce an egocentric update (Klatzky, Loomis, Beall, Chance, & Golledge, 1998).
- Participants were not familiar with the content of the VEs (the city was fictitious) and had only a small amount of time with no repetitions, nor an option to stop and study a feature a little longer if they desired. These all may have limited the knowledge acquisition compared to what one would naturally do in the real world. Furthermore, as stated by Montello et al. (2004), the field of view and the allocation of attention can be confounding to the acquisition of route knowledge. Future experiments may take these issues into consideration to better establish the ecological validity of the findings from this research.

User characteristics

- The experiment was a between-subject experiment (older vs. younger adults). This was a necessary condition, because a longitudinal study is beyond the limits of a Ph.D. project. While the participants were counterbalanced for expertise and gender, there may be "cohort effect" differences that affect the outcome. Specifically, it is plausible that experience with technology and 3D content and learning from media was different for 20-30-year group and the 65-75-year-old group. The fact that younger people may have been more exposed to virtual settings may have affected their encoding of information differently (Brown, 2000) compared to older people who, due to their lack of exposure, may have had one more barrier to encoding (Czaja & Lee, 2007). This is a hard problem to solve, even if one was willing to conduct a longitudinal study, as technology tends to change quickly over one's lifetime. The findings should be viewed with some caution from this perspective.
- Spatial abilities and memory capacity were not used as inclusion criteria for participation. These two factors were included for control purposes, and treated at the analysis stage (thus as post-hoc experimental hypotheses). Even though the distribution of people with higher and lower abilities was reasonable for a post-hoc analysis, i.e., there was a "naturally counterbalanced" distribution; a future approach may be to ensure this at the recruitment stage.

6. Conclusions and outlook

6.1 Scientific contributions

This thesis examined how the three main dimensions of (i) *visualization type*, (ii) *use context*, and (iii) *user characteristics* interrelate in the recall of information in a route learning task with the use of 3D virtual environments. The findings from the thesis clearly demonstrate that the recall of visuospatial information with 3D visualizations is **significantly influenced by all of the dimensions** cited above: Participants' route recall performance is affected by the **levels of *realism*** in the VEs, **the type of visuospatial information** to be recalled **in short- and long-term**, as well as the **participants' *age*,** and ***cognitive abilities***. Thus, this thesis provides further evidence that these dimensions are all important to consider when studying route learning. A general conclusion thus is that 3D visualizations used for route learning and possibly other learning tasks should be designed based on a combination of these primary dimensions. Given the above, this thesis contributes to GIScience, geovisualization, and the cognitive design of visualizations in two ways:

- ❖ Understanding the effects of **visual realism**—as presented in VEs used for route learning—on **human memory,** and by extension, on human cognition in general, specifically **for people of different ages** and with different spatial abilities and memory capacity**.**
- ❖ Proposing **empirically validated visualization design guidelines for younger and older people** for the efficient encoding and decoding of visuospatial information **during route learning,** not only **in the short-term** but **also in the long-term.**

These visualization design guidelines are summarized in the following section.

6.2 Key findings and visualization design guidelines

This section summarizes all of the findings discussed previously and provides two tables that are meant as generalized guidelines. The guidelines are also applicable to other VE navigational learning scenarios, in connection to the use of visual realism in route learning based on the "lessons learned" in this Ph.D. project regarding (i) characteristics of the *route recall task* and *stage* and (ii) characteristics of the *users*.

3D visualizations with different levels of realism	
Recall task & recall stage	**Design guidelines**
- All tasks - Both recall stages	***Use photorealism selectively. Beware of user preferences.*** ❖ VEs similar to the Mixed VE can be used overall for route learning training successfully. ❖ VEs similar to the Realistic VE likely will not be overall better for navigational learning than those that are similar to the Abstract VE. If the goal is different (e.g., scene gist recognition, a virtual reality field visit, exploration of remote locations), however, photorealism may be fine. ❖ Majority of users may prefer fully photorealistic VEs (such as the Realistic VE in this thesis) if they have no prior experience of using the VE for the task ahead of them. Leaving the choice of VE to the users may be a bad idea in some cases.
- Turn-by-turn visuospatial task - Both recall stages	***Think about the information type: Does solving the task require use of both visual and spatial memory? Beware of "confident" users.*** ❖ VEs similar to the Mixed VE are likely to be better than the Abstract and Realistic VEs for inherently visuospatial tasks where perspective remains the same. ❖ VEs similar to the Realistic VE rather than the Abstract VE can be ok to use for visual tasks in the short-term, while long-term recall should be treated with caution. ❖ Be cautious about confident users. Using VEs similar to the Mixed VE might moderate overconfidence.
- Map-sketching visuospatial task - Both recall stages	***Think about the cognitive processes needed for the tasks: Do the users have to perform mental rotations? Consider giving people active tasks if the goal is learning.*** ❖ If the task is predominantly spatial (i.e., visual details do not matter), VEs similar to the Mixed and Abstract VEs can be used. ❖ Actively engaging users (such as making them draw a sketch) might support information retention. If the goal is long-term learning, enable users to execute and rehearse with active tasks.
- Visual task - Both recall stages	***If the task mostly relies on photographic information, photorealism can be ok but the selective use of photorealism is even better.*** ❖ VEs in such as the Mixed VE should be used rather than the Abstract and Realistic VE types. Realistic VE types are better than the Abstract VEs, the complete lack of photorealistic cues is not helpful in visual tasks. ❖ For long-term recall, VEs such as the Mixed and Realistic VE can be used. However, additional measures might be a good idea (e.g., provide more training trials).
- Spatial task - Immediate recall stage	***If the task relies mostly on spatial memory, adding photorealistic information does not improve or impair performance.*** ❖ All VE types from this thesis have the same route recall outcomes for spatial tasks. With some caution, you can decide which level of realism to use based on other criteria.

3D visualizations with different levels of realism	
User characteristics	**Design guidelines**
Age	***For an inclusive design, remember the age-related cognitive decline and its effects. Balance the amount of information to lighten the cognitive load of older adults. Pay attention to potential issues of metacognition (overconfidence).*** ❖ In this thesis, on average, effects of cognitive decline were observable in older adults' performance. Older adults were also overconfident in their response accuracy. ❖ In similar contexts, VEs similar to the Mixed VE can be used as an "exemplary design solution" to amplify memory performance for both age groups for short- and long-term recall. ❖ VEs similar to the Mixed VE might moderate overconfidence to healthier levels. ❖ Older adults may struggle more with some task types than the younger. Specifically, if the task involves perspective changing from egocentric to allocentric, they might need support. ❖ Older adults may have metacognition issues somewhat more severe than younger adults. This affects not only their confidence, but also their preferences. Since they did not change their preference to match their performance in this thesis as much as the younger adults, we assume more of the older adults were unable to recognize which tool was helping them more. Making tool recommendations to older adults may require different strategies.
Spatial abilities	***Take differences in spatial abilities into account.*** ❖ Spatial abilities, as measured based on mental rotation abilities, do matter in overall route learning performance. ❖ In this thesis, there was no interaction between visualization type and spatial abilities. This suggests that the VEs such as the Mixed VE might be helpful to *all* people, regardless of their mental rotation abilities. ❖ Older people, even those with higher spatial abilities, may have an upper limit in long-term recall. Assume that spatial abilities matter but higher spatial abilities may not help the older people as much as the younger.
Visuospatial memory capacity	***Take visuospatial memory abilities into account.*** ❖ Visuospatial memory capacity of the individuals can affect their route learning performance. ❖ VEs similar to the Mixed VE can be helpful to *all* people also irrespective of their visuospatial memory capacity. ❖ The VSM test is a good test for predicting route recall overall. However, it seems like people's VSM scores do not correlate with their success in tasks that are predominantly visual. For this task type, perhaps another test would predict performance better.

6.3 Outlook

There are many new directions for further exploration and research that can confirm the findings and expand the research conducted within the framework of this thesis. Here, these suggestions are once again structured around the three dimensions investigated, many of which are linked to the limitations listed earlier.

Visualization type: The following is an expansion on the design suggestions this thesis investigated:

- ❖ The Mixed VE resulted from the joint investigation of two parameters: Using *realistic highlighting* at *critical*-for-navigation locations. In an effort to separate the two parameters, future research could attempt to understand the implications that realistic highlighting in *random locations* presents for route learning.
- ❖ Extending the above idea from the reverse perspective, as *visual realism* is only one type of highlighting mechanism; The value of introducing visual realism is discussed throughout this thesis. However, other mechanisms can prove to be relevant, especially in relation to visual variables such as the use of colors and symbols.
- ❖ Instead of counterbalancing the textures, a future project could expand their investigation and focus on the "memorability" and "visual saliency" of textures as main variables to further confirm their full effects.
- ❖ Perspective-switching visuospatial task (map-sketching) required a switch in perspective, and this proved to be difficult for participants, especially those with lower abilities. This issue could be potentially averted with the simultaneous representation of a 2D bird's eye (aerial) view of the visual scene while encoding the route from a first-person perspective. The effect of the additional 2D representation may assist people in better understanding the spatial configuration of the area they are walking through, providing new opportunities for better acquisition of survey knowledge.

Use context: This thesis has focused on route recall in a VE of a fictitious city. Possible changes in the use context are as follows:

- ❖ Expanding the use context from an urban to a rural setting (Brosset, Claramunt, & Saux, 2008) where the density of the environment is different and the navigational landmarks are potentially differentiated in shape and size, differently facilitating route recall.
- ❖ Introducing active involvement throughout the encoding stage (e.g., participants virtually walking or driving a route, interacting with the scene instead of passively observing a drive-through) may result in different—and potentially higher—recall outcomes that may highlight differences across the visualization types more prominently.
- ❖ A real-world evaluation of the value that the Mixed VE design has for route learning. By having participants take a route in the actual world, after having learnt it in the Mixed VE, we can evaluate the transferability of the learning process.
- ❖ Transforming this study into a longitudinal route training study, to evaluate longer-term effects on additional cognitive abilities, supporting brain training for healthy aging (Cassarino & Setti, 2015; Millington, 2012; Walton et al., 2014).

User characteristics: An important contribution of this thesis was the investigation of group differences, especially based on age, and the nuanced implications tied to it. Further focus on the users can be along the following lines:

- ❖ By conducting a longitudinal study, the younger participants can be re-tested at several points in the future and confounds linked to "cohort effects" may be examined. Additionally, potential declines in spatial abilities and memory capacity can be better linked to aging.

- By pre-screening people for their spatial abilities and memory capacity, especially in studies examining aging and spatial cognition, one may clearly separate the two variables. Such a screening may result in a more representative sample of people with typically higher and lower spatial abilities and memory capacities for *each* age group. Additionally, measuring cognitive load explicitly, and examining individual recall strategies may shed more light on the amount of information people are capable of storing (Norman, 1976).
- The fact that naïve realism as a concept seems to be partially confirmed in this thesis (especially for the older age group), highlights the need to expand visualization explorations to achieve a better alignment between user preference and user recall performance.

Methodology: The design of an experiment can impact the quality of the results. Ensuring that the measured variable is adequately isolated from confounding variables guarantees clarity in the findings. Ideas for further improvement include:

- With technological advancements, the use of more immersive VR settings may potentially improve encoding and enhance visual quality. Additionally, they could affect route recall accuracy for users with different characteristics (e.g., age and abilities).
- Introduction of active involvement of the user in the encoding of information could affect the performance differences regarding *recall stage*s and could benefit from further examination.
- Introduction of repeated trials could affect recall accuracy and confidence (calibration errors). Recall accuracy may increase with repetition, and metacognition may increase too, potentially facilitating healthier confidence levels.

References

Aginsky, V., Harris, C., Rensink, R., & Beusmans, J. (1997). Two strategies for learning a route in a driving simulator. *Journal of Environmental Psychology*, *17*(4), 317–331. https://doi.org/10.1006/jevp.1997.0070

Anders, T. R., Fozard, J. L., & Lillyquist, T. D. (1972). Effects of age upon retrieval from short-term memory. *Developmental Psychology*, *6*(2), 214–217. https://doi.org/10.1037/h0032103

Angel, L., Fay, S., Bouazzaoui, B., Baudouin, A., & Isingrini, M. (2010). Protective role of educational level on episodic memory aging: An event-related potential study. *Brain and Cognition*, *74*(3), 312–323. https://doi.org/10.1016/j.bandc.2010.08.012

Annett, M. (1992). Spatial ability in subgroups of left-and right-handers. *British Journal of Psychology*, *83*(4), 493–515. https://doi.org/10.1111/j.2044-8295.1992.tb02455.x

Arias-Hernandez, R., Green, T. M., & Fisher, B. (2012). From cognitive amplifiers to cognitive prostheses: Understandings of the material basis of cognition in visual analytics. *Interdisciplinary Science Reviews*, *37*(1), 4–18. https://doi.org/10.1179/0308018812Z.0000000001

Bak, T. H., Nissan, J. J., Allerhand, M. M., & Deary, I. J. (2014). Does bilingualism influence cognitive aging? *Annals of Neurology*, *75*(6), 959–963. https://doi.org/10.1002/ana.24158

Barbeau, E., Didic, M., Tramoni, E., Felician, O., Joubert, S., Sontheimer, A., … Poncet, M. (2004). Evaluation of visual recognition memory in MCI patients. *Neurology*, *62*(8), 1317–1322. https://doi.org/10.1212/01.WNL.0000120548.24298.DB

Barrash, J. (1994). Age-Related Decline in Route Learning Ability. *Developmental Neuropsychology*, *10*(3), 189–201. https://doi.org/10.1080/87565649409540578

Bertin, J. (1983). *Semiology of Graphics: Diagrams, Networks, Maps. Semiology of Graphics: Diagrams, Networks, Maps.* Madison: University of Wisconsin Press. https://doi.org/10.1037/023518

Betrancourt, M. (2005). The Animation and Interactivity Principles in Multimedia Learning. In *The Cambridge Handbook of Multimedia Learning* (pp. 287–296). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.019

Bétrancourt, M., & Tversky, B. (2000). Effect of computer animation on users' performance: A review. *Travail Humain : A Bilingual and Multi-Disciplinary Journal in Human Factors*, *63*(4), 311–329.

Bleisch, S., & Nebiker, S. (2008). Connected 2D and 3D visualizations for the interactive exploration of spatial information. In *Proc. of 21th ISPRS Congress, Beijing, China* (pp. 1037–1042).

Boer, A., Coltekin, A., & Clarke, K. C. (2013). Evaluating Web-Based Geovisualizations Online: A Case Study with Abstraction-Realism Spectrum In Focus. In *ICC2013*. Dresden. Retrieved from http://icaci.org/files/documents/ICC_proceedings/ICC2013/_extendedAbstract/284_proceeding.pdf

Boér, A., Çöltekin, A., & Clarke, K. C. (2013). An evaluation of web-based geovisualizations for different levels of abstraction and realism – What do users predict? In *International Cartograhic Conference, ICC 2013, Dresden, Germany*.

Borkin, M. A., Gajos, K. Z., Peters, A., Mitsouras, D., Melchionna, S., Rybicki, F. J., … Pfister, H. (2011). Evaluation of Artery Visualizations for Heart Disease Diagnosis. *IEEE Transactions on Visualization and Computer Graphics*, *17*(12), 2479–2488. https://doi.org/10.1109/TVCG.2011.192

Borkin, M. A., Vo, A. A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., & Pfister, H. (2013). What makes a visualization memorable. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2306–2315. https://doi.org/10.1109/TVCG.2013.234

Boukhelifa, N., Bezerianos, A., Isenberg, T., & Fekete, J. D. (2012). Evaluating sketchiness as a visual variable for the depiction of qualitative uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2769–2778. https://doi.org/10.1109/TVCG.2012.220

Bowman, D. A., & Mcmahan, R. P. (2007). Virtual Reality: How Much Immersion Is Enough?

Computer, *40*(7), 36–43. Retrieved from http://ezlibproxy.unisa.edu.au/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=25758828&site=ehost-live

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Brewer, C. A. (1996). Guidelines for selecting colors for diverging schemes on maps. *Cartographic Journal*, *33*(2), 79–86. https://doi.org/10.1179/caj.1996.33.2.79

Brewer, C. A. (2015). *Designing better Maps: A Guide for GIS users*. ESRI press.

Brosset, D., Claramunt, C., & Saux, E. (2008). Wayfinding in natural and urban environments: A comparative study. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *43*(1), 21–30. https://doi.org/10.3138/carto.43.1.21

Brown, A. M., Lindsey, D. T., & Guckes, K. M. (2011). Color names, color categories, and color-cued visual search: Sometimes, color perception is not categorical. *Journal of Vision*, *11*(12). https://doi.org/10.1167/11.12.1

Brown, J. S. (2000). Growing Up: Digital: How the Web Changes Work, Education, and the Ways People Learn. *Change: The Magazine of Higher Learning*, *32*(2), 11–20. https://doi.org/10.1080/00091380009601719

Brügger, A., Fabrikant, S. I., & Çöltekin, A. (2017). An Empirical Evaluation of Three Elevation Change Symbolization Methods along Routes in Bicycle Maps. *Cartography and Geographic Information Science*, *44*(5), 436–451. https://doi.org/10.1080/15230406.2016.1193766

Brügger, A., Richter, K.-F., & Fabrikant, S. I. (2019). How does navigation system behavior influence human behavior? *Cognitive Research: Principles and Implications*, *4*(1), 5. https://doi.org/10.1186/s41235-019-0156-5

Brychtova, A., & Coltekin, A. (2016). An Empirical User Study for Measuring the Influence of Colour Distance and Font Size in Map Reading Using Eye Tracking. *Cartographic Journal*, *53*(3), 202–212. https://doi.org/10.1179/1743277414Y.0000000103

Burigat, S., & Chittaro, L. (2007). Navigation in 3D virtual environments: Effects of user experience and location-pointing navigation aids. *International Journal of Human Computer Studies*, *65*(11), 945–958. https://doi.org/10.1016/j.ijhcs.2007.07.003

Burnett, S. A., Lane, D. M., & Dratt, L. M. (1982). Spatial ability and handedness. *Intelligence*, *6*(1), 57–68. https://doi.org/10.1016/0160-2896(82)90020-4

Byrne, J. H. (2017). *Learning and Memory: A Comprehensive Reference*. *Learning and Memory: A Comprehensive Reference*. Academic Press. https://doi.org/10.1016/c2015-1-01759-8

Card, S., Mackinlay, J., & Shneiderman, B. (1999). *Readings in Information Visualisation. Using Vision to Think.: Using Vision to Think*. *Morgan Kaufmann*.

Carpendale, M. S. T. (2003). Considering Visual Variables as a basis for Information Visualisation. *Cartographica: International Journal for Geographic Information and Geovisualization*.

Cassarino, M., & Setti, A. (2015). Environment as "Brain Training": A review of geographical and physical environmental influences on cognitive ageing. *Ageing Research Reviews*, *23*, 167–182. https://doi.org/10.1016/j.arr.2015.06.003

Claramunt, C., & Winter, S. (2007). Structural salience of elements of the city. *Environment and Planning B: Planning and Design*, *34*(6), 1030–1050. https://doi.org/10.1068/b32099

Cockburn, A., & McKenzie, B. (2002). Evaluating the Effectiveness of Spatial Memory in 2D and 3D Physical and Virtual Environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 203–210). https://doi.org/10.1145/503376.503413

Çöltekin, A., Lokka, I. E., & Zahner, M. (2016). On the usability and usefulness of 3D (Geo)Visualizations - A focus on virtual reality environments. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*. https://doi.org/10.5194/isprsarchives-XLI-B2-387-2016

Çöltekin, Arzu. (2019). What contributes to the complexity of visuospatial displays? Abstraction,

Scale and Perception. In *International Cartographic Association Joint Commission Workshop*.
Çöltekin, Arzu, Bleisch, S., Andrienko, G., & Dykes, J. (2017). Persistent challenges in geovisualization – a community perspective. *International Journal of Cartography*, 115–139. https://doi.org/10.1080/23729333.2017.1302910
Çöltekin, Arzu, Francelet, R., Richter, K. F., Thoresen, J., & Fabrikant, S. I. (2018). The effects of visual realism, spatial abilities, and competition on performance in map-based route learning in men. *Cartography and Geographic Information Science*, *45*(4), 339–353. https://doi.org/10.1080/15230406.2017.1344569
Çöltekin, Arzu, Heil, B., Garlandini, S., & Fabrikant, S. I. (2009). Evaluating the Effectiveness of Interactive Map Interface Designs: A Case Study Integrating Usability Metrics with Eye-Movement Analysis. *Cartography and Geographic Information Science*, *36*(1), 5–17. https://doi.org/10.1559/152304009787340197
Coluccia, E., & Louse, G. (2004). Gender differences in spatial orientation: A review. *Journal of Environmental Psychology*, *24*(3), 329–340. https://doi.org/10.1016/j.jenvp.2004.08.006
Cover, S. A., Ezquerra, N., O'Brien, J. F., Rowe, R., Gadacz, T., & Palm, E. (1993). Interactively deformable models for surgery simulation. *IEEE Computer Graphics and Applications*, *13*(6), 68–75. https://doi.org/10.1109/38.252559
Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, *19*(1), 51–57. https://doi.org/10.1177/0963721409359277
Credé, S., Thrash, T., Hölscher, C., & Fabrikant, S. I. (2020). The advantage of globally visible landmarks for spatial learning. *Journal of Environmental Psychology*, *67*, 101369. https://doi.org/10.1016/j.jenvp.2019.101369
Czaja, S. J., & Lee, C. C. (2007). The impact of aging on access to technology. *Universal Access in the Information Society*, *5*(4), 341. https://doi.org/10.1007/s10209-006-0060-x
Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: a tool for unwelding visuo–spatial memory. *Neuropsychologia*, *37*(10), 1189–1199. https://doi.org/10.1016/S0028-3932(98)00159-6
Dibiase, D., Maceachren, A. M., Krygier, J. B., & Reeves, C. (1992). Animation and the role of map design in scientific visualization. *Cartography & Geographic Information Systems*, *19*(4), 201–214. https://doi.org/10.1559/152304092783721295
Dodson, C. S., Bawa, S., & Krueger, L. E. (2007). Aging, metamemory, and high-confidence errors: A misrecollection account. *Psychology and Aging*, *22*(1), 122. https://doi.org/10.1037/0882-7974.22.1.122
Dodson, C. S., Bawa, S., & Slotnick, S. D. (2007). Aging, source memory, and misrecollections. *Journal of Experimental Psychology: Learning Memory and Cognition*, *33*(1), 169. https://doi.org/10.1037/0278-7393.33.1.169
Dodson, C. S., Koutstaal, W., & Schacter, D. L. (2000). Escape from illusion: Reducing false memories. *Trends in Cognitive Sciences*, *4*(10), 391–397. https://doi.org/10.1016/S1364-6613(00)01534-5
Dodson, C. S., & Krueger, L. E. (2006). I misremember it well: Why older adults are unreliable eyewitnesses. *Psychonomic Bulletin and Review*, *13*(5), 770–775. https://doi.org/10.3758/BF03193995
Dubuc, B. (2002). Memory and learning. Retrieved January 4, 2020, from http://thebrain.mcgill.ca/flash/d/d_07/d_07_p/d_07_p_tra/d_07_p_tra.html
Dykes, J., MacEachren, A. M., & Kraak, M.-J. (2005). *Exploring Geovisualization*. Elsevier.
Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). Manual for kit of factor-referenced cognitive tests. *Princeton, NJ: Educational Testing Service*, (August).
Evans, G. W., & Pezdek, K. (1980). Cognitive mapping: Knowledge of real-world distance and location information. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(1), 13. https://doi.org/10.1037/0278-7393.6.1.13

Fabrikant, S. I. S., & Lobben, A. (2009). Introduction: Cognitive Issues in Geographic Information Visualization. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *44*(3), 139–143. https://doi.org/10.3138/carto.44.3.139

Fabrikant, S. I., Maggi, S., & Montello, D. R. (2014). 3D network spatialization: Does it add depth to 2D representations of semantic proximity? *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *8728*, 34–47. https://doi.org/10.1007/978-3-319-11593-1_3

Fillekes, M. P., Giannouli, E., Kim, E. K., Zijlstra, W., & Weibel, R. (2019). Towards a comprehensive set of GPS-based indicators reflecting the multidimensional nature of daily mobility for applications in health and aging research. *International Journal of Health Geographics*, *18*(1), 17. https://doi.org/10.1186/s12942-019-0181-0

Fillekes, M. P., Kim, E. K., Trumpf, R., Zijlstra, W., Giannouli, E., & Weibel, R. (2019). Assessing older adults' daily mobility: A comparison of GPS-derived and self-reported mobility indicators. *Sensors*, *19*(20), 4551. https://doi.org/10.3390/s19204551

Finlayson, N. J., Zhang, X., & Golomb, J. D. (2017). Differential patterns of 2D location versus depth decoding along the visual hierarchy. *NeuroImage*, *147*, 507–516. https://doi.org/10.1016/j.neuroimage.2016.12.039

Foos, P. W. (1989). Adult age differences in working memory. *Psychology and Aging*, *4*(3), 269–275. https://doi.org/10.1037//0882-7974.4.4.500

Foreman, N. (2010). Virtual Reality in Psychology. *THEMES IN SCIENCE AND TECHNOLOGY EDUCATION*, *2*(1–2), 225–252.

Fozard, J. L., Rietsema, J., Bouma, H., & Graafmans, J. A. M. (2000). Gerontechnology: Creating enabling environments for the challenges and opportunities of aging. *Educational Gerontology*, *26*(4), 331–334. https://doi.org/10.1080/036012700407820

Fricke, M., & Bock, O. (2018). Egocentric navigation is age-resistant: First direct behavioral evidence. *Current Neurobiology*, *9*(2), 69–75.

Gale, N., Golledge, R. G., Pellegrino, J. W., & Doherty, S. (1990). The acquisition and integration of route knowledge in an unfamiliar neighborhood. *Journal of Environmental Psychology*, *10*(1), 3–25. https://doi.org/10.1016/S0272-4944(05)80021-0

Gallagher, S. M., & Keenan, M. (2009). Stimulus equivalence and the Mini Mental Status Examination in the elderly. *European Journal of Behavior Analysis*, *10*(2), 159–165. https://doi.org/10.1080/15021149.2009.11434316

Garlandini, S., & Fabrikant, S. I. (2009). Evaluating the Effectiveness and Efficiency of Visual Variables for Geographic Information Visualization. In *International Conference on Spatial Information Theory* (pp. 195–211). https://doi.org/10.1007/978-3-642-03832-7_12

Golledge, R. G., Dougherty, V., & Bell, S. (1995). Acquiring Spatial Knowledge: Survey Versus Route-Based Knowledge in Unfamiliar Environments. *Annals of the Association of American Geographers*, *85*(1), 134–158. https://doi.org/10.1111/j.1467-8306.1995.tb01798.x

Gonzalez, C. (1996). Does animation in user interfaces improve decision making? In *Proceedings of the SIGCHI conference on Human factors in computing systems common ground - CHI '96* (pp. 27–34). New York, New York, USA: ACM Press. https://doi.org/10.1145/238386.238396

Goodman, J., Brewster, S., & Gray, P. (2004). Older people, mobile devices and navigation. *HCI and the Older Population*, 13–14.

Gras, D., Gyselinck, V., Perrussel, M., Orriols, E., & Piolino, P. (2013). The role of working memory components and visuospatial abilities in route learning within a virtual environment. *Journal of Cognitive Psychology*, *25*(1), 38–50. https://doi.org/10.1080/20445911.2012.739154

Griffin, A. L., & Fabrikant, S. I. (2012). More maps, more users, more devices means more cartographic challenges. *Cartographic Journal*, *49*(4), 298–301. https://doi.org/10.1179/0008704112Z.00000000049

Grön, G., Wunderlich, A. P., Spitzer, M., Tomczak, R., & Riepe, M. W. (2000). Brain activation during human navigation: Gender-different neural networks as substrate of performance. *Nature*

Neuroscience, *3*(4), 404–408. https://doi.org/10.1038/73980
Gross, A. L., Parisi, J. M., Spira, A. P., Kueider, A. M., Ko, J. Y., Saczynski, J. S., … Rebok, G. W. (2012). Memory training interventions for older adults: A meta-analysis. *Aging and Mental Health*, *16*(6), 722–734. https://doi.org/10.1080/13607863.2012.667783
Gyselinck, V., Meneghetti, C., Bormetti, M., Orriols, E., Piolino, P., & De Beni, R. (2013). Considering spatial ability in virtual route learning in early aging. *Cognitive Processing*, *14*(3), 309–316. https://doi.org/10.1007/s10339-013-0557-1
Head, D., & Isom, M. (2010). Age effects on wayfinding and route learning skills. *Behavioural Brain Research*, *209*(1), 49–58. https://doi.org/10.1016/j.bbr.2010.01.012
Heft, H. (1979). The role of environmental features in route-learning: Two exploratory studies of way-finding. *Environmental Psychology and Nonverbal Behavior*, *3*(3), 172–185. https://doi.org/10.1007/BF01142591
Hegarty, M., Smallman, H. S., Stull, A. T., & Canham, M. S. (2009). Naïve Cartography: How Intuitions about Display Configuration Can Hurt Performance. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *44*(3), 171–186. https://doi.org/10.3138/carto.44.3.171
Herbert, G., & Chen, X. (2015). A comparison of usefulness of 2D and 3D representations of urban planning. *Cartography and Geographic Information Science*, *42*(1), 22–32. https://doi.org/10.1080/15230406.2014.987694
Herman, J. F., & Coyne, A. C. (1980). Mental manipulation of spatial information in young and elderly adults. *Developmental Psychology*, *16*(5), 537. https://doi.org/10.1037/0012-1649.16.5.537
Hertzog, C., & Rypma, B. (1991). Age differences in components of mental-rotation task performance. *Bulletin of the Psychonomic Society*, *29*(3), 209–212. https://doi.org/10.3758/BF03342680
Hodge, G. (2008). *The geography of aging: Preparing communities for the surge in seniors*. *The Geography of Aging: Preparing Communities for the Surge in Seniors*. McGill-Queen's Press-MQUP.
Hort, J., Laczó, J., Vyhnálek, M., Bojar, M., Bureš, J., & Vlček, K. (2007). Spatial navigation deficit in amnestic mild cognitive impairment. *Proceedings of the National Academy of Sciences*, *104*(10), 4042–4047. https://doi.org/10.1073/pnas.0611314104
Huk, T. (2006). Who benefits from learning with 3D models? the case of spatial ability. *Journal of Computer Assisted Learning*, *22*(6), 392–404. https://doi.org/10.1111/j.1365-2729.2006.00180.x
Iachini, T., Iavarone, A., Senese, V. P., Ruotolo, F., & Ruggiero, G. (2009). Visuospatial Memory in Healthy Elderly, AD and MCI: A Review. *Current Aging Science*, *2*(1), 43–59. https://doi.org/10.2174/1874612810902010043
Iaria, G., Palermo, L., Committeri, G., & Barton, J. J. S. (2009). Age differences in the formation and use of cognitive maps. *Behavioural Brain Research*, *196*(2), 187–191. https://doi.org/10.1016/j.bbr.2008.08.040
Inagaki, H., Meguro, K., Shimada, M., Ishizaki, J., Okuzumi, H., & Yamadori, A. (2002). Discrepancy between mental rotation and perspective-taking abilities in normal aging assessed by Piaget's Three-Mountain Task. *Journal of Clinical and Experimental Neuropsychology*, *24*(1), 18–25. https://doi.org/10.1076/jcen.24.1.18.969
Isaacson, M., Wahl, H. W., Shoval, N., Oswald, F., & Auslander, G. (2016). The relationship between spatial activity and wellbeing-related data among healthy older adults: An exploratory geographic and psychological analysis. In *Cross-Cultural and Cross-Disciplinary Perspectives in Social Gerontology* (pp. 203–219). Springer, Singapore. https://doi.org/10.1007/978-981-10-1654-7_11
Ishikawa, T., & Montello, D. R. (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology*, *52*(2), 93–129.

https://doi.org/10.1016/j.cogpsych.2005.08.003
Isola, P., Xiao, J., Parikh, D., Torralba, A., & Oliva, A. (2014). What makes a photograph memorable? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(7), 1469–1482. https://doi.org/10.1109/TPAMI.2013.200
Isola, P., Xiao, J., Torralba, A., & Oliva, A. (2011). What makes an image memorable? *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 145–152. Retrieved from http://cvcl.mit.edu/memorableImages.html
Jameson, A. (2008). Adaptive Interfaces and Agents. In *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (pp. 105–130). https://doi.org/10.1201/9781420088861.ch6
Karimpur, H., Röser, F., & Hamburger, K. (2016). Finding the Return Path: Landmark Position Effects and the Influence of Perspective. *Frontiers in Psychology*, *7*, 1956. https://doi.org/10.3389/fpsyg.2016.01956
Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language*, *48*(4), 704–721. https://doi.org/10.1016/S0749-596X(02)00504-1
Kessels, R. P. C., Meulenbroek, O., Fernandez, G., & Olde Rikkert, M. G. M. (2010). Spatial working memory in aging and mild cognitive impairment: Effects of task load and contextual cueing. *Aging, Neuropsychology, and Cognition*, *17*(5), 556–574. https://doi.org/10.1080/13825585.2010.481354
Kinateder, M., Ronchi, E., Nilsson, D., Kobes, M., Müller, M., Pauli, P., & Mühlberger, A. (2014). Virtual reality for fire evacuation research. In *2014 Federated Conference on Computer Science and Information Systems, IEEE* (pp. 313–321). https://doi.org/10.15439/2014F94
Kirasic, K. C. (1991). Spatial cognition and behavior in young and elderly adults: implications for learning new environments. *Psychology and Aging*, *6*(1), 10–18. https://doi.org/10.1037//0882-7974.6.1.10
Klatzky, R. L. (1998). Allocentric and Egocentric Spatial Representations: Definitions, Distinctions, and Interconnections. In *In: Freksa C., Habel C., Wender K.F. (eds) Spatial Cognition. Lecture Notes in Computer Science, vol 1404. Springer, Berlin, Heidelberg*. https://doi.org/10.1007/3-540-69342-4_1
Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science*, *9*(4), 293–298. https://doi.org/10.1111/1467-9280.00058
Klencklen, G., Després, O., & Dufour, A. (2012). What do we know about aging and spatial cognition? Reviews and perspectives. *Ageing Research Reviews*, *11*(1), 123–135. https://doi.org/10.1016/J.ARR.2011.10.001
Klippel, A., Richter, K.-F., & Hansen, S. (2005). Structural Salience as a Landmark. *Workshop Mobile Maps*.
Klippel, A., & Winter, S. (2005). Structural salience of landmarks for route directions. *Cohn A.G., Mark D.M. (Eds) Spatial Information Theory. COSIT 2005. Lecture Notes in Computer Science*, *3693*, 347–362. https://doi.org/10.1007/11556114_22
Koutstaal, W., & Schacter, D. L. (1997). Gist-Based False Recognition of Pictures in Older and Younger Adults. *Journal of Memory and Language*, *37*(4), 555–583.
Krejtz, K., Çöltekin, A., Duchowski, A. T., & Niedzielska, A. (2017). Using coeffcient K to distinguish ambient/focal visual attention during cartographic tasks. *Journal of Eye Movement Research*, *10*(2), 3. https://doi.org/https://doi.org/10.16910/jemr.10.2.3
Lee, S. A., & Spelke, E. S. (2010). Two systems of spatial representation underlying navigation. *Experimental Brain Research*, *206*(2), 179–188. https://doi.org/10.1007/s00221-010-2349-5
Lenneberg, E. H. (1961). Color Naming, Color Recognition, Color Discrimination: A Re-Appraisal. *Perceptual and Motor Skills*, *12*(3), 375–382. https://doi.org/10.2466/pms.1961.12.3.375
Liao, H., Dong, W., Peng, C., & Liu, H. (2017). Exploring differences of visual attention in pedestrian

navigation when using 2D maps and 3D geo-browsers. *Cartography and Geographic Information Science*, *44*(6), 474–490. https://doi.org/10.1080/15230406.2016.1174886
Lloyd, R. E., & Bunch, R. L. (2005). Individual differences in map reading spatial abilities using perceptual and memory processes. *Cartography and Geographic Information Science*, *32*(1), 33–46. https://doi.org/10.1559/1523040053270774
Logie, R. H. (1986). Visuo-spatial Processing in Working Memory. *The Quarterly Journal Of Experimental Psychology*, *38*(2), 229–247. https://doi.org/10.1080/14640748608401596
Lokka, I. E., & Çöltekin, A. (2017). Remembering What we See : Designing Virtual Environments to Improve Visuo-Spatial Recall for Navigation Tasks. In *Proceedings of the 28th International Cartographic Conference (ICC2017)*.
Lokka, I. E., & Çöltekin, A. (2019). Toward optimizing the design of virtual environments for route learning: empirically assessing the effects of changing levels of realism on memory. *International Journal of Digital Earth*, *12*(2), 137–155. https://doi.org/10.1080/17538947.2017.1349842
Lokka, I. E., & Çöltekin, A. (2020). Perspective switch and spatial knowledge acquisition: effects of age, mental rotation ability and visuospatial memory capacity on route learning in virtual environments with different levels of realism. *Cartography and Geographic Information Science*, *47*(1), 14–27. https://doi.org/10.1080/15230406.2019.1595151
Lokka, I. E., Çöltekin, A., Wiener, J., Fabrikant, S. I., & Röcke, C. (2018). Virtual environments as memory training devices in navigational tasks for older adults. *Scientific Reports*, *8*(1), 10809. https://doi.org/10.1038/s41598-018-29029-x
Luebke, D., Reddy, M., Cohen, J. D., Varshney, A., Watson, B., & Huebner, R. (2003). Terrain Level of Detail. In *Level of Detail for 3D Graphics*. Morgan Kaufmann. https://doi.org/10.1016/b978-155860838-2/50009-1
MacDorman, K. F. (2006). Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley. *ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science*, 26–29. https://doi.org/10.1093/scan/nsr025
MacEachren, A. M. (2004). *How Maps Work: Representation, Visualization, and Design*. *How Maps Work: Representation, Visualization, and Design*. https://doi.org/10.14714/CP24.757
Maggi, S., Fabrikant, S. I., Imbert, J.-P., & Hurter, C. (2016). How Do Display Design and User Characteristics Matter in Animations? An Empirical Study with Air Traffic Control Displays. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *51*(1), 25–37. https://doi.org/10.3138/cart.3176
Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: A structural MRI and neuropsychological analysis. *Hippocampus*, *16*(12), 1091–1101. https://doi.org/10.1002/hipo.20233
Maquet, P. (2001). The role of sleep in learning and memory. *Science*, *294*(5544), 1048–1052. https://doi.org/10.1126/science.1062856
Mastin, L. (2010). Short-term (working) memory. Retrieved December 28, 2019, from https://human-memory.net/short-term-working-memory/
McIntire, J. P., Ellis, S. A., Harrington, L. K., & Havig, P. R. (2014). Subjective evaluations of multiple three-dimensional displays by a stereo-deficient viewer: an interesting case study. *Display Technologies and Applications for Defense, Security, and Avionics VIII; and Head- and Helmet-Mounted Displays XIX*, *9086*, 908605. https://doi.org/10.1117/12.2054051
McIntire, J. P., Havig, P. R., & Geiselman, E. E. (2014). Stereoscopic 3D displays and human performance: A comprehensive review. *Displays*, *35*(1), 18–26. https://doi.org/10.1016/j.displa.2013.10.004
McLeod, S. (2010). Long-Term Memory. Retrieved January 12, 2020, from https://www.simplypsychology.org/long-term-memory.html
McLeod, S. (2017). Multi Store Model of Memory. Retrieved December 28, 2019, from https://www.simplypsychology.org/multi-store.html

McMaster, R. B., & Shea, K. S. (1992). Generalization in digital cartography. *Washington, DC: Association of American Geographers*.
Mcshane, R., Gedling, K., Keene, J., Fairburn, C., Jacoby, R., & Hope, T. (1998). Getting lost in dementia: A longitudinal study of a behavioral symptom. *International Psychogeriatrics*, *10*(3), 253–260. https://doi.org/10.1017/S1041610298005365
Meijer, F., Geudeke, B. L., & Van Den Broek, E. L. (2009). Navigating through virtual environments: Visual realism improves spatial cognition. *Cyberpsychology and Behavior*, *12*(5), 517–521. https://doi.org/10.1089/cpb.2009.0053
Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *2*(1), 1–21. https://doi.org/10.1016/S0022-5371(63)80063-8
Meneghetti, C., Borella, E., Muffato, V., Pazzaglia, F., & De Beni, R. (2014). Environment learning from spatial descriptions: The role of perspective and spatial abilities in young and older adults. *Freksa C., Nebel B., Hegarty M., Barkowsky T. (Eds) Spatial Cognition IX. Spatial Cognition 2014. Lecture Notes in Computer Science*, *8684*. https://doi.org/10.1007/978-3-319-11215-2_3
Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158
Millington, B. (2012). Use it or lose it: Ageing and the politics of brain training. *Leisure Studies*, *31*(4), 429–446. https://doi.org/10.1080/02614367.2011.589865
Modigliani, V., & Seamon, J. G. (1974). Transfer of information from short- to long-term memory. *Journal of Experimental Psychology*, *102*(5), 768–772. https://doi.org/10.1037/h0036374
Moffat, S. D., & Resnick, S. M. (2002a). Effects of age on virtual environment place navigation and allocentric cognitive mapping. *Behavioral Neuroscience*, *116*(5), 851–859. https://doi.org/10.1037/0735-7044.116.5.851
Moffat, S. D., & Resnick, S. M. (2002b). Effects of age on virtual environment place navigation and allocentric cognitive mapping. *Behavioral Neuroscience*. https://doi.org/10.1037/0735-7044.116.5.851
Moffat, S. D., Zonderman, A. B., & Resnick, S. M. (2001). Age differences in spatial memory in a virtual environment navigation task. *Neurobiology of Aging*, *22*(5), 787–796. https://doi.org/10.1016/S0197-4580(01)00251-2
Moffat, S., Zonderman, A. B., & Resnick, S. M. (2001). Age differences in spatial memory in a virtual environment navigation task. *Neurobiology of Aging*, *22*(5), 787–796. https://doi.org/10.1016/S0197-4580(01)00251-2
Montello, D. R. (1998). A new framework for understanding thr acquisition of spatial knowledge in large-scale environments. In *Spatial and Temporal Reasoning in Geographic Information Systems* (pp. 143–154). https://doi.org/10.1088/1748-6041/6/2/025001
Montello, D. R., Waller, D., Hegarty, M., & Richardson, A. E. (2004). Spatial Memory of Real Environments, Virtual Environments, and Maps. *Human Spatial Memory: Remembering Where*, 251–285.
Montello, D., & Sas, C. (2006). Human Factors of Wayfinding in Navigation. In *International Encyclopedia of Ergonomics and Human Factors, Second Edition - 3 Volume Set*. CRC Press. https://doi.org/10.1201/9780849375477.ch394
Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley. *IEEE Robotics & Automation Magazine*, *19*(2), 98–100. https://doi.org/10.1109/MRA.2012.2192811
Muffato, V., Della Giustina, M., Meneghetti, C., & De Beni, R. (2015). Age-related differences in pointing accuracy in familiar and unfamiliar environments. *Cognitive Processing*, *16*(1), 313–317. https://doi.org/10.1007/s10339-015-0720-y
Muffato, V., Meneghetti, C., Di Ruocco, V., & De Beni, R. (2017). When young and older adults learn a map: The influence of individual visuo-spatial factors. *Learning and Individual Differences*, *53*, 114–121. https://doi.org/10.1016/j.lindif.2016.12.002

Nemmi, F., Boccia, M., & Guariglia, C. (2017). Does aging affect the formation of new topographical memories? Evidence from an extensive spatial training. *Aging, Neuropsychology, and Cognition*, *24*(1), 29–44. https://doi.org/10.1080/13825585.2016.1167162

Norman, D. A. (1976). *Memory and attention: An introduction to human information processing, 2nd ed. John Wiley & Sons*.

O'Malley, M., Innes, A., & Wiener, J. M. (2018). How do we get there? Effects of cognitive aging on route memory. *Memory and Cognition*, *46*(2), 274–284. https://doi.org/10.3758/s13421-017-0763-7

Olson, J. M. (1997). Multimedia in Geography: Good, Bad, Ugly, or Cool? *Annals of the Association of American Geographers*, *87*(4), 571–578. https://doi.org/10.1111/1467-8306.00068

Özgen, E. (2004). Language, learning, and color perception. *Current Directions in Psychological Science*, *13*(3), 95–98. https://doi.org/10.1111/j.0963-7214.2004.00282.x

Pardilla-Delgado, E., & Payne, J. D. (2017). The Deese-Roediger-Mcdermott (DRM) task: A simple cognitive paradigm to investigate false memories in the laboratory. *Journal of Visualized Experiments*, (119), 54793. https://doi.org/10.3791/54793

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, *17*(2), 299–320. https://doi.org/10.1037/0882-7974.17.2.299

Partala, T., Nurminen, A., Vainio, T., Laaksonen, J., Laine, M., & Väänänen, J. (2010). Salience of visual cues in 3D city maps. In *Proceedings of the 24th British Computer Society Conference on Human-Computer Interaction* (pp. 428–432). https://doi.org/10.14236/ewic/hci2010.51

Peters, S., Jahnke, M., Murphy, C. E., Meng, L., & Abdul-Rahman, A. (2017). Cartographic enrichment of 3D city models—state of the art and research perspectives. *Lecture Notes in Geoinformation and Cartography,* 207-230. https://doi.org/10.1007/978-3-319-25691-7_12

Plesa, M. A., & Cartwright, W. (2008). Evaluating the Effectiveness of Non-Realistic 3D Maps for Navigation with Mobile Devices. In *Meng L., Zipf A., Winter S. (eds) Map-based Mobile Services. Lecture Notes in Geoinformation and Cartography* (pp. 80–104). https://doi.org/10.1007/978-3-540-37110-6_5

Plouffe, L., & Kalache, A. (2010). Towards global Age-Friendly cities: Determining urban features that promote active aging. *Journal of Urban Health*, *87*(5), 733–739. https://doi.org/10.1007/s11524-010-9466-0

Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, *81*(1), 10. https://doi.org/10.1037/h0027470

Raubal, M., & Winter, S. (2002). Enriching wayfinding instructions with local landmarks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *2478*. https://doi.org/10.1007/3-540-45799-2_17

Rendell, P. G., Castel, A. D., & Craik, F. I. M. (2005). Memory for proper names in old age: a disproportionate impairment? *The Quarterly Journal of Experimental Psychology. Section A*, *58*(1), 54–71. https://doi.org/10.1080/02724980443000188

Richardson, A. E., Montello, D. R., & Hegarty, M. (1999). Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory and Cognition*, *27*(4), 741–750. https://doi.org/10.3758/BF03211566

Richter, K.-F., & Winter, S. (2014). *Landmarks*. Springer Cham Heidelberg New York Dordrecht London. https://doi.org/10.1007/978-3-319-05732-3

Risden, K., Czerwinski, M. P., Munzner, T., & Cook, D. B. (2000). An initial examination of ease of use for 2D and 3D information visualizations of web content. *International Journal of Human Computer Studies*, *53*(5), 695–714. https://doi.org/10.1006/ijhc.2000.0413

Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision*, *7*(2), 17. https://doi.org/10.1167/7.2.17

Röser, F., Hamburger, K., Krumnack, A., & Knauff, M. (2012). The structural salience of landmarks: Results from an on-line study and a virtual environment experiment. *Journal of Spatial Science*,

57(1), 37–50. https://doi.org/10.1080/14498596.2012.686362
Röser, F., Krumnack, A., Hamburger, K., & Knauff, M. (2012). A four factor model of landmark salience - A new approach. In *Proceedings of the 11th International Conference on Cognitive Modeling, ICCM 2012* (pp. 82–87).
Roth, R. E., Çöltekin, A., Delazari, L., Filho, H. F., Griffin, A., Hall, A., … van Elzakker, C. P. J. M. (2017). User studies in cartography: opportunities for empirical research on interactive maps and visualizations. *International Journal of Cartography*, *3*, 61–89. https://doi.org/10.1080/23729333.2017.1288534
Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, *13*(4), 140–144. https://doi.org/10.1111/j.0963-7214.2004.00293.x
Sanders, B., Wilson, J. R., & Vandenberg, S. G. (1982). Handedness and Spatial Ability. *Cortex*, *18*(1), 79–89. https://doi.org/10.1016/S0010-9452(82)80020-8
Schofield, C. P., & Honoré, S. (2009). Generation Y and learning. *Designing High-Impact Learning: Translating Research into Practice.*
Seipel, S. (2013). Evaluating 2D and 3D geovisualisations for basic spatial assessment. *Behaviour & Information Technology*, *32*(8), 845–858. https://doi.org/10.1080/0144929X.2012.661555
Semmo, A., Trapp, M., Kyprianidis, J. E., & Döllner, J. (2012). Interactive Visualization of Generalized Virtual 3D City Models using Level-of-Abstraction Transitions. *Computer Graphics Forum*, *31*, 885–894. https://doi.org/10.1111/j.1467-8659.2012.03081.x
Shepherd, I. D. H. (2008). Travails in the Third Dimension: A Critical Evaluation of Three-Dimensional Geographical Visualization. https://doi.org/10.1002/9780470987643.ch10
Slocum, T. A., McMaster, R., Kessler, F., & Howard, H. (1999). *Thematic cartography and visualization*. *Methods*.
Smallman, H. S., & John, M. (2005a). Naive Realism: Limits of Realism as a Display Principle. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *49*(17), 1564–1568. https://doi.org/10.1177/154193120504901714
Smallman, H. S., & John, M. S. (2005b). Naive realism: misplaced faith in realistic displays. *Ergonomics in Design: The Quarterly of Human Factors Applications*, *13*(3), 6–13. https://doi.org/10.1177/106480460501300303
Smallman, Harvey S., & Cook, M. B. (2011). Naive Realism: Folk fallacies in the design and use of visual displays. *Topics in Cognitive Science*, *3*(3), 579–608. https://doi.org/10.1111/j.1756-8765.2010.01114.x
Sorrows, M. E., & Hirtle, S. C. (1999). The nature of landmarks for real and electronic spaces. In *International Conference on Spatial Information Theory* (pp. 37–50). https://doi.org/10.1007/3-540-48384-5_3
Stankiewicz, B. J., & Kalia, A. A. (2007). Acquisition of Structural Versus Object Landmark Knowledge. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(2), 378–390. https://doi.org/10.1037/0096-1523.33.2.378
Steichen, B., Ashman, H., & Wade, V. (2012). A comparative survey of Personalised Information Retrieval and Adaptive Hypermedia techniques. *Information Processing and Management*, *48*(4), 698–724. https://doi.org/10.1016/j.ipm.2011.12.004
Steichen, B., & Fu, B. (2019). Towards Adaptive Information Visualization - A Study of Information Visualization Aids and the Role of User Cognitive Style. *Frontiers in Artificial Intelligence*, *2*, 22. https://doi.org/10.3389/frai.2019.00022
Strickrodt, M., O'Malley, M., & Wiener, J. M. (2015). This place looks familiar-how navigators distinguish places with ambiguous landmark objects when learning novel routes. *Frontiers in Psychology*, *6*, 1936. https://doi.org/10.3389/fpsyg.2015.01936
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7
Taylor, H. A., & Tversky, B. (1996). Perspective in Spatial Descriptions. *Journal of Memory and Language*, *35*(3), 371–391. https://doi.org/10.1006/jmla.1996.0021

Terry, W. S. (2016). *Learning and Memory - Basic Principles, Processes, and Procedures*. *IEEE Transactions on Information Theory*. https://doi.org/10.1177/0363546505279922
Thorndyke, P., & Hayes-Roth, B. (1982). Differences in Spatial Knowledge Acquired and Navigation from Maps. *Cognitive Psychology*, *14*, 560–589. https://doi.org/10.1016/0010-0285(82)90019-6
Tombaugh, T. N., & McIntyre, N. J. (1992). The Mini-Mental State Examination: A Comprehensive Review. *Journal of the American Geriatrics Society*, *40*(9), 922–935. https://doi.org/10.1111/j.1532-5415.1992.tb01992.x
Tory, M. (2003). Mental Registration of 2D and 3D Visualizations (An Empirical Study). *Proceedings of the IEEE Visualization Conference*, 371–378. https://doi.org/10.1109/VISUAL.2003.1250396
Tufte, E. (2001). *The visual display of quantitative information, second edition. Technometrics*.
Tulving, E. (1972). Episodic and Semantic Memory. In *Organization of memory* (pp. 381–403). Retrieved from http://alumni.media.mit.edu/~jorkin/generals/papers/Tulving_memory.pdf
Turcotte, M. (2013). Family caregiving: What are the consequences? *Statistics Canada*. https://doi.org/Catalogue no. 75-006-X
van der Ham, I. J. M., & Claessen, M. H. G. (2020). How age relates to spatial navigation performance: functional and methodological considerations. *Ageing Research Reviews*. https://doi.org/10.1016/j.arr.2020.101020
Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, *47*(2), 599–604. https://doi.org/10.2466/pms.1978.47.2.599
Waller, D., & Greenauer, N. (2007). The role of body-based sensory information in the acquisition of enduring spatial representations. *Psychological Research*, *71*(3), 322–332. https://doi.org/10.1007/s00426-006-0087-x
Waller, D., & Lippa, Y. (2007). Landmarks as beacons and associative cues: Their role in route learning. *Memory & Cognition*, *35*(5), 910–924. https://doi.org/http://dx.doi.org/10.3758/BF03193465
Walton, C. C., Mowszowski, L., Lewis, S. J. G., & Naismith, S. L. (2014). Stuck in the mud: Time for change in the implementation of cognitive training research in ageing? *Frontiers in Aging Neuroscience*, *6*(MAR), 2012–2015. https://doi.org/10.3389/fnagi.2014.00043
Ware, C. (2004). *Information Visualization: Perception for Design.* San Francisco, CA, USA: Morgan Kaufmann.
Wiener, J. M., de Condappa, O., Harris, M. a, & Wolbers, T. (2013). Maladaptive bias for extrahippocampal navigation strategies in aging humans. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(14), 6012–6017. https://doi.org/10.1523/JNEUROSCI.0717-12.2013
Wilkening, J., & Fabrikant, S. I. (2013). How users interact with a 3D geo-browser under time pressure. *Cartography and Geographic Information Science*, *40*(1), 40–52. https://doi.org/10.1080/15230406.2013.762140
Winter, S. (2003). Route adaptive selection of salient features. *Kuhn W., Worboys M.F., Timpf S. (Eds) Spatial Information Theory. Foundations of Geographic Information Science. COSIT 2003. Lecture Notes in Computer Science*, *2825*. https://doi.org/10.1007/978-3-540-39923-0_23
Wolbers, T., & Hegarty, M. (2010). What determines our navigational abilities? *Trends in Cognitive Sciences*, *14*(3), 138–146. https://doi.org/10.1016/j.tics.2010.01.001
Wolfe, J. M. (1998). Visual memory: What do you know about what you saw? *Current Biology*, *8*(9), 303–304. https://doi.org/10.1016/s0960-9822(98)70192-7
Wood, J., Kirschenbauer, S., Döllner, J., Lopes, A., & Bodum, L. (2005). Using 3D in Visualization. In *Exploring Geovisualization* (pp. 293–312). Elsevier.
World Health Organization. (2011). *National Institute on Aging, National Institutes of Health, U.S. Department of Health and Human Services*. *Global Health and Aging*.
Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-

formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482. https://doi.org/10.1002/cne.920180503

Zanola, S., Fabrikant, S. I., & Çöltekin, A. (2009). The effect of realism on the confidence in spatial data quality in stereoscopic 3D displays. In *24th International Cartographic Conference (ICC 2009)* 2, 1–9.

Zelinski, E. M., & Light, L. L. (1988). Young and older adults' use of context in spatial memory. *Psychology and Aging*, *3*(1), 99. https://doi.org/10.1037/0882-7974.3.1.99

Zhong, J. Y., & Moffat, S. D. (2016). Age-related differences in associative learning of landmarks and heading directions in a virtual navigation task. *Frontiers in Aging Neuroscience*, *8*(122), 1–11. https://doi.org/10.3389/fnagi.2016.00122

II. PUBLICATIONS

Publication I

Lokka, I.E., & Çöltekin, A. (2019). Toward optimizing the design of virtual environments for route learning: empirically assessing the effects of changing levels of realism on memory. *International Journal of Digital Earth*, 12(2), 137-155. https://doi.org/10.1080/17538947.2017.1349842

International Journal of Digital Earth

ISSN: 1753-8947 (Print) 1753-8955 (Online) Journal homepage: https://www.tandfonline.com/loi/tjde20

Toward optimizing the design of virtual environments for route learning: empirically assessing the effects of changing levels of realism on memory

Ismini E. Lokka & Arzu Çöltekin

To cite this article: Ismini E. Lokka & Arzu Çöltekin (2019) Toward optimizing the design of virtual environments for route learning: empirically assessing the effects of changing levels of realism on memory, International Journal of Digital Earth, 12:2, 137-155, DOI: 10.1080/17538947.2017.1349842

To link to this article: https://doi.org/10.1080/17538947.2017.1349842

Published online: 19 Jul 2017.

INTERNATIONAL JOURNAL OF DIGITAL EARTH
2019, VOL. 12, NO. 2, 137–155
https://doi.org/10.1080/17538947.2017.1349842

Check for updates

Toward optimizing the design of virtual environments for route learning: empirically assessing the effects of changing levels of realism on memory

Ismini E. Lokka and Arzu Çöltekin

Department of Geography, University of Zurich, Zurich, Switzerland

ABSTRACT

Broadly, this paper is about designing memorable 3D geovisualizations for spatial knowledge acquisition during (virtual) navigation. Navigation is a fundamentally important task, and even though most people navigate every day, many find it difficult in unfamiliar environments. When people get lost in an unfamiliar environment, or are unable to remember a route that they took, they might feel anxiety, disappointment and frustration; and in real world, such incidents can be costly, and at times, life-threatening. Therefore, in this paper, we study the design decisions in terms of visual realism in a city model, propose a visualization design optimized for route learning, implement and empirically evaluate this design. The evaluation features a navigational route learning task, where we measure short- and long-term recall accuracy of 42 participants with varying spatial abilities and memory capacity. Our findings provide unique empirical evidence on how design choices affect memory in route learning with geovirtual environments, contributing toward empirically verified design guidelines for digital cities.

ARTICLE HISTORY
Received 30 April 2017
Accepted 29 June 2017

KEYWORDS
3D representation; digital city; digital earth; virtual reality; visualization

1. Introduction

A *virtual reality* (VR) idea has fascinated people for decades, and early instances of VR were created in 1960s (Sutherland 1965). Because a virtual world can be used as a 'spatial lab', VR also found an audience in geography (e.g. Fisher and Unwin 2001). A peak in the excitement for potential contributions of an all-encompassing geographical VR to education and exploration led to the proposition of a *Digital Earth* (Gore 1998). The term *virtual environments* (VE) extends the VR concept into a visualization environment that can *also* feature simulated or fictional worlds. VEs in geography (geovirtual environments or GeoVEs) were suggested as a research priority in GIScience nearly two decades ago (MacEachren et al. 1999; MacEachren and Kraak 2001; Slocum et al. 2001), because they, in Slocum et al.'s (2001, 62) words, 'fundamentally change our traditional way of acquiring spatial knowledge'. In the past two decades, impressive progress has been made in technology, promising 'better' GeoVEs. However, we still know very little on how the visualization design in a GeoVE affects spatial knowledge acquisition. This paper contributes toward a better understanding of how (and how much) various elements of design, especially levels of realism, contribute to the recall effectiveness of GeoVEs as learning environments.

CONTACT Ismini E. Lokka ismini-eleni.lokka@geo.uzh.ch

2. Related work

Below we provide a review of the related work on: (i) cognitive processes during navigation involving memory (ii) visualization design considerations and (iii) the individual differences in cognitive abilities relevant to (real or virtual) navigation tasks.

2.1. Cognitive processes related to navigation: the indispensable role of memory

Spatial cognition research on navigation largely reports on attention (i.e. what do people notice), and information encoding (i.e. what kind of mental notes they take) during navigation. In such studies, an important factor for route learning appears to be the *perspective* from which people experience the route. It has been proposed that an egocentric perspective during learning leads to the so-called *route knowledge*, that is, a 'procedure' of necessary movements to reach a point (Gillner and Mallot 1998), whereas an allocentric perspective leads to a 'global' understanding of the surroundings, termed *survey knowledge* (Lobben 2004). This position is debated, however, irrespective of its validity or whether it is survey or route knowledge, *memory* plays key role in all stages of spatial learning related to navigational tasks.

Memory is a multifaceted cognitive process. First of all, different *memory types* are involved in acquiring spatial knowledge. It is not straightforward to assign route- or survey-knowledge acquisition into one of the common memory systems (e.g. implicit/explicit) (Montello et al. 2004). Nonetheless, classifications have been proposed depending on the type of information one must recall. One such classification of memory types, relevant to this paper, refers to *visual*, *spatial* and *visuospatial* information. Although the visual and spatial memories are tightly coupled in some tasks (Klauer and Zhao 2004), we adopt the position that there are distinct memory systems that encode/store and decode/retrieve visual and spatial information (Della Sala et al. 1999); and the two often 'cooperate' (i.e. visuospatial). Notably, during the decoding, there are subtle differences in the processes, for example, the terms *recall* and *recognition* are distinguished (Freund, Brelsford, and Atkinson 1969). We use the term 'recall' for the memory tasks used in this paper for the sake of simplicity.

Memory systems are also classified based on duration, most commonly as *short-* and *long-term*. An event is stored in the short-term memory almost instantly, arguably for a few seconds (Peterson and Peterson 1959). Short-term memory, especially the 'few-seconds' definition, is often used interchangeably with the term *working memory*, although there are arguments for distinguishing the two. The most common argument is that the working memory does not *store* the information at all, while short-term memory stores it for a short time (Cowan 2008). The capacity of the working memory is limited to four to seven objects (Miller 1956; Cowan 2001), and the amount of detail stored regarding these objects is quite limited (Luck and Hollingworth 2008). Short-term memories are transient, whereas long-term memories are often reinforced with rehearsal, and once transferred to the 'long-term storage', they are assumed to have an infinite duration (Luck and Hollingworth 2008). We use the term *short-term memory* for recall rates several minutes after the experience (different than what is considered working memory), and long-term memory for knowledge decoding roughly after an hour or longer.

2.2. Visualization design considerations for route learning in VEs

Realistic and abstract geovisualizations are both used as learning aids in various contexts, and are important in route learning (Çöltekin et al. 2017). Realistic VEs are popular in testing navigational tasks, as they allow for a safe environment and more experimental control than the real world studies (Loomis and Blascovich 1999; Dünser et al. 2006; Bülthoff, Campos, and Meilinger 2008). In such contexts, it has been consistently shown with other types of geovisualizations that the visualization type and design affect performance with a variety of spatial tasks (Bleisch and Dykes 2014; Roth et al.

2017). Even subtle differences in visual variables (Garlandini and Fabrikant 2009), such as color (Brychtová and Çöltekin 2017), shading (Bernabé Poveda, Angel, and Çöltekin 2015; Biland and Çöltekin 2017), symbology type (Brügger, Fabrikant, and Çöltekin 2016) or levels of realism (Wilkening and Fabrikant 2011) can affect how well people execute various spatial tasks. While there are some considerations in comparing 2D and 3D (Cockburn and McKenzie 2002; Çöltekin, Lokka, and Zahner 2016), studies on how to design a GeoVE to make route learning more effective are scarce.

A key decision regarding visualization design appears to be about the *amount of information*, that is, too much information can increase cognitive load and impair performance with spatial tasks (Smallman and John 2005; Plesa and Cartwright 2008; Hegarty, Smallman, and Stull 2012; Dong and Liao 2016; Liao et al. 2016). VEs are often designed as photorealistically as possible, with the objective to replicate the real world and increase immersion (even though immersion does not necessarily require photorealism, see McMahan 2003). In this paper, we ask if 'too much information' can impair performance in spatial tasks, is photorealism a threat to GeoVEs' effective use in certain contexts? At this point, we do not have clear guidelines on how much realism should be included in GeoVEs.

Abstract visualizations (ideally) remove task-irrelevant information, and guide users' attention to the relevant information for a specific task (Scheiter et al. 2009), and have been shown to be more effective than realistic visualizations in some spatial tasks (Hegarty, Canham, and Fabrikant 2010; Wilkening and Fabrikant 2011). In support of abstraction, Sanchez and Branaghan (2009) demonstrated that adding more detail on a display affects map reading negatively, and impairs recall success in a route learning task. Conversely, a highly realistic visualization might have higher *ecological validity* than an abstract alternative, given that a VE simulates the real world (Kattenbeck 2015). Besides, a realistic VE includes readily *recognizable* elements, which might support memory (Christou and Bülthoff 1999; Meijer, Geudeke, and van den Broek 2009; Borkin et al. 2013). The realistic looking visual elements that people can name might be better retained in memory compared to more abstract shapes and structures, because of the so-called dual channel assumption; that is, people utilize two cognitive channels (e.g. verbal and visual) simultaneously (Mayer and Moreno 2003).

Some efforts to manage the *level of detail* (LOD) in VEs focus on presenting features with different LODs; using 'more detail' selectively as highlighting mechanisms, for example, in focus + context visualizations (Betrancourt 2005; Semmo et al. 2012; Peters et al. 2017), and using additional objects as landmarks (Parush and Berman 2004). Other efforts focus on the technical aspects of defining and creating LOD (e.g. https://www.citygml.org/), or managing LOD by removing perceptually irrelevant details (e.g. Bektaş and Çöltekin 2011).

Besides the amount, the semantic *quality* of the information (*what* is shown) can influence route learning performance in a VE. For example, landmarks play a significant role in spatial knowledge acquisition (Richter and Winter 2014). Landmark is a difficult term to define, however, *structural*, *visual* and *semantic* saliency are important characteristics for landmarks (Raubal and Winter 2002; Klippel and Winter 2005). For structural salience, the impact of *location* appears to be important (e.g. Röser et al. 2012). Röser et al. (2012) found that the landmarks at the decision points (intersections) are the most important, especially those at the direction of the turn. Visual salience is also important in the context of navigational learning, as attention is critical in memory and learning (Itti, Koch, and Niebur 1998). Besides landmarks, Lynch (1960) identifies paths (routes) to be 'predominant elements in [the observer's] image' (Lynch 1960, 47), and Claramunt and Winter (2007) posit that street networks are cognitively (semantically) salient. We believe that for a memorable GeoVE, all three aspects of saliency (visual/structural/semantic) must be considered.

An interesting additional aspect in visual realism studies is that seemingly people's intuitive preferences do not always match their performance with realistic visualizations. Two theories have been proposed in relation to this mismatch between performance and preference: Smallman and John's (2005) *naive realism* theory suggests an unfounded preference toward realism, which was later followed by *naive cartography* in which the effect was reproduced for enhanced displays with animations and 3D (Hegarty et al. 2009). These theories provide an interesting insight into how our visualization-related choices could be misguided, and should be considered in studies such as ours.

2.3. Individual and group differences

People differ in learning with visualizations based on various abilities, age, expertise and other factors in their background (Slocum et al. 2001). For example, Huk (2006) demonstrates that people with higher spatial abilities (high-spatial) benefit more from 3D in learning than people with lower spatial abilities (low-spatial). Spatial abilities that are most relevant in navigational tasks are proposed to be: (mentally) visualizing objects, relating objects, mental rotation, path integration and spatial updating (Richter and Winter 2014). Standardized psychometric tests (Ekstrom et al. 1976) can predict people's effectiveness in using visualizations (Hegarty and Waller 2004). Spatial abilities, as measured by standardized tests, can have a significant influence on people's performance also in navigation tasks (Schinazi et al. 2013). It is interesting to note that the spatial abilities might play a role even in naive realism. In Smallman and Cook's (2011) study, all participants preferred the realistic displays *before* the experiment, but only high-spatial participants adjusted their preference to abstract displays *after*; suggesting that low-spatial participants struggle assessing self-performance.

Various other factors in a user's background, such as experience, age (Salthouse 2006) or gender (Parush and Berman 2004) might also affect route learning performance. In the scope of this paper, we analyze how spatial abilities and memory capacity interact with route learning performance, and counterbalance for other factors that might affect performance in route learning.

3. Hypotheses

Based on the previous work cited above, we propose a VE that is designed with specific *amount* and *type* of information presented in key locations, that is, we use photo-textures only for selected parts of the VE. These parts are thus 'highlighted' and should act as anchoring points or landmarks. With our proposed virtual world (MixedVE), route recall should be easier than with a RealisticVE, or an AbstractVE with no textures. We specifically hypothesize that:

- Participants' visual, spatial and visuospatial recall performance will be best with the MixedVE, irrespective of their spatial abilities, both in the short- and long-term
- Participant's overall recall performance with the RealisticVE will be better than with the AbstractVE, as the RealisticVE provides more visual cues
- High-spatial participants will overall perform better with the RealisticVE, and specifically with tasks that are more demanding on the memory than the low-spatial participants.

4. Experimental design

In a mixed factorial design ($3 \times 2 \times 4$), we tested the three levels of realism as our *independent variables*: (i) the AbstractVE with no photo-textures (baseline), (ii) the RealisticVE (fully photorealistic) and (iii) the MixedVE designed based on previous knowledge on levels of realism and landmark theories (Figure 1) (Lokka and Çöltekin 2016, 2017). Throughout the manuscript, we call these VEs *visualization types*. Four different task types (Visual, Spatial, Visuospatial and Map/perspective switch), and individual differences based on two criteria (spatial ability, memory capacity) are considered as potentially moderating factors. Note that we study the recall rates (i) right after the route learning task (short-term memory: *Stage1*), (ii) about an hour later (long-term memory: *Stage2*) and (iii) a week later (long-term memory: *Stage3*). Thus, we examine if (potential) differences in memory performance with the three VEs would persist.

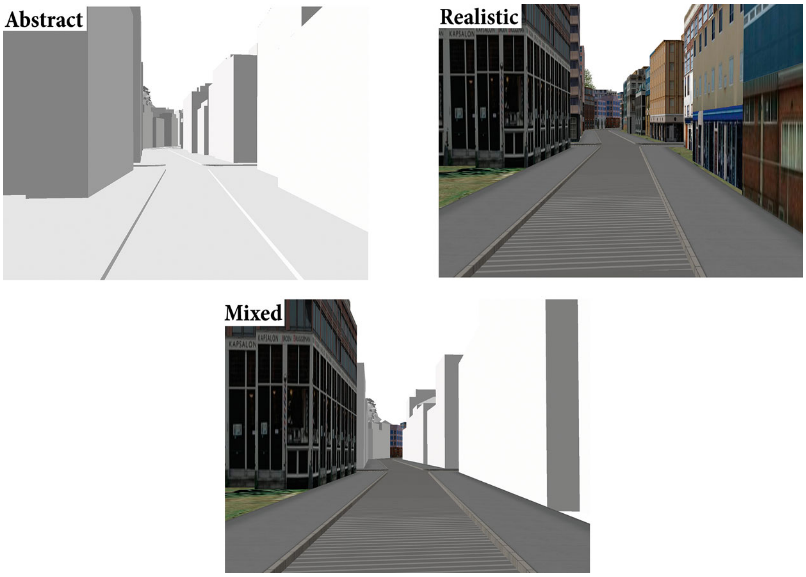


Figure 1. Screenshots illustrating the three VEs (not to scale).

As *dependent variables*, we report on *recall accuracy* for all visualization types and task types, and participants' visualization *preferences* before and after the experiment.

4.1. Participants

Forty-two participants (M = 27 years, 23 women) voluntarily took part in the experiment based on informed consent. The age range was kept to 20–30, because aging affects memory (Park et al. 2002). All participants were university students (undergraduate to PhD) in different degree programs and were recruited through individual contact. We measured their *spatial abilities* using a Mental Rotation Task (MRT, Vandenberg and Kuse 1978), and *visuospatial memory capacities* using a Visuospatial Memory Test (VSM, Ekstrom et al. 1976).

4.2. Materials

4.2.1. Apparatus

We performed the experiment in controlled lab, where we back-projected the VEs as videos on a large screen (230 × 140 cm), which was 2.2 m away from the participant (Figure 2). We used an off-the-shelf experimental software to deliver all visualizations and tasks.

Stimuli. All VEs represented the same fictitious city, which was created using procedural modeling. We kept the lighting conditions constant, the buildings similar in size and in architectural style, trees and intersections with comparable visual and spatial characteristics. From each VE, we created fly-bys of two pre-selected routes as videos. All videos were shown only once at the same eye-level, the same scale, extent and speed, simulating a drive (duration: 100 s, speed: 30 km/h). The AbstractVE was rendered in grayscale without photo-textures (Figure 1, top-left). The MixedVE had photo-textures on selected buildings at the turn points toward the direction of the turn, and the road network was photo-textured to highlight the spatial structure (Figure 1, bottom). The contents of the photo-textures were counterbalanced with regards to visual saliency (i.e. using visual-saliency algorithms by Itti et al. 1998) and memorability (e.g. Borkin et al. 2013; Lokka and Çöltekin 2017) in the MixedVE. The RealisticVE was fully photo-textured (Figure 1, top-right).

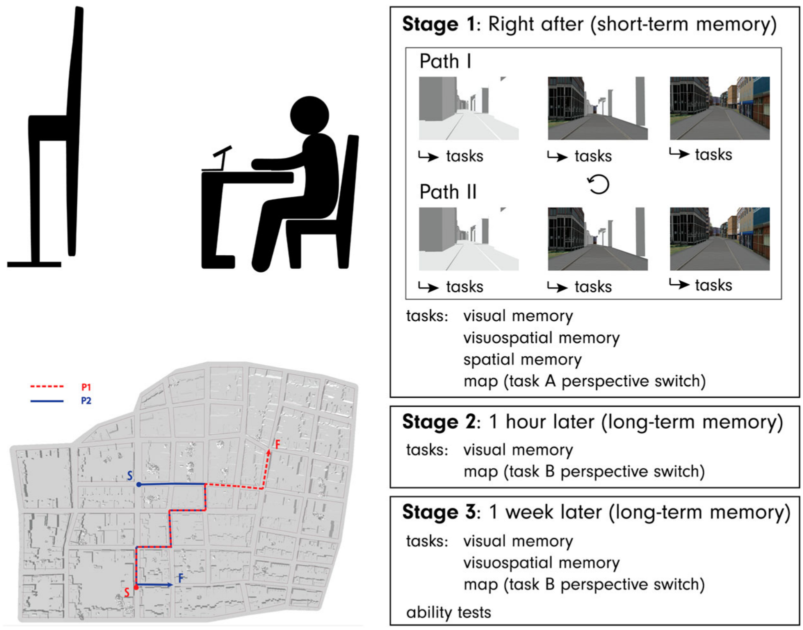


Figure 2. Experimental setup (left-top), the two routes (left-bottom) and the procedure (right).

We prepared two routes; each consisted of seven intersections (three turns toward the left, three turns toward the right and one continuing straight, as presented in Figure 2).

4.3. Tasks

The participants were instructed to *memorize a route* from a starting to an ending point as they watched the videos in a wayfinding scenario. For each visualization type, they responded to a set of questions (in Stages 1, 2 and 3) which we categorize into four task types:

Visual memory (VM) tasks: Based on six screenshots from each VE (three correct, three false), participants' task was to identify whether they had *seen the image or not*. They answered using a 6-point Likert scale ranging from 'definitely-seen' to 'definitely-not-seen'. This task type was used in all experimental stages.

Spatial memory (SM) tasks: In this set, participants were asked to identify the direction they were facing at the end of the route (starting orientation was given), and the number of turns they took during the virtual drive. These two questions were asked only in Stage 1. They were left out from Stages 2 and 3, as it would be impossible to distinguish from which visualization type they recalled the information after having watched all videos.

Visuospatial memory (VS) tasks: Participants marked which direction they turned at all seven intersections one-by-one, based on screenshots, which appeared in the same order and perspective as in the VEs. Additionally, they were asked to identify the start- and end-points of their route from four options (only one was correct). These questions were asked in Stages 1 and 3. We excluded the VS tasks in Stage 2 because of time limits.

MapTask (MT)/perspective switch: This task type requires a perspective switch (from egocentric to allocentric), and can be seen as a special instance of the VS tasks, they are predominantly spatial, but some visual cues were also provided ('aerial' view screenshots from each VE). Participants were to first identify (Stage 1, MapTaskA), then actively reproduce (Stages 2–3, MapTaskB) the route based on a top-down 2D view. In Stage 1, four options were provided with one correct answer (MapTaskA), and in Stages 2–3, participants drew sketches (start- and end-points were marked) on paper (MapTaskB).

4.4. Procedure

Upon arrival, we welcomed the participants, and they read and signed the consent form. Right after, participants stated their *preference* between the three VEs (shown as screenshots) for a hypothetical route learning task. Then the main experiment began. Participants watched the three VEs for the two pre-selected routes (thus, six videos) in a randomized order. After each video, they answered a set of questions with all four task types. After the first three videos and associated questions, participants took a small break (to counter learning and fatigue). After viewing all six videos and solving associated tasks, Stage 1 was completed. Stage 2 followed with two task types (Visual & MapTaskB) regarding all six videos shown in Stage 1 and stated their preference *again* between the three visualizations. The duration of the experiment was on average 1 h:30 m for Stages 1–2. Participants came back 6–8 days later for Stage 3, responded to a demographic questionnaire, and continued with three task types (Visual, Visuospatial and MapTaskB), after which we conducted the MRT and VSM tests. Stage 3 lasted approximately 1 h. An overview of the procedure is shown in Figure 2.

5. Results

Below, we provide participants' overall recall accuracy with the VEs, followed by how different task types interact with recall accuracy. Then, we examine how participants' spatial abilities (based on the MRT) and memory capacity (based on the VSM) interact with their recall accuracy with each VE and task type. We then demonstrate the long-term recall rates based on a comparison between the three stages for comparable tasks. Furthermore, we report on participants' preferences regarding the tested visualization types *before* and *after* the experiment.

The recall accuracy was calculated as the proportion of correct answers to all answers. For the MapTaskB, we counted the errors in number of turns, the number of left/right turns, the sequence and the direction for the start- and end-points. Statistical analyses were conducted using *R* with $\alpha = .05$. We report associated p-values <.05 as statistically significant, and mark the p-values that fall between [.1–.5] as statistical trends. We include estimations of effect size (η_p^2), for which .01 is considered small, .06 medium, and .14 and above, large (Ellis 2010).

5.1. Short-term memory: Stage 1

Figure 3 demonstrates that for the short-term memory tasks, participants' recall accuracy is highest with the MixedVE. The MixedVE improves recall accuracy by roughly 12.1% in comparison to the AbstractVE, and 7.7% in comparison to the RealisticVE. Both differences are statistically significant with a large effect size (Table 1, 'overall'). We also see that the participants' overall recall accuracy is higher with the RealisticVE than with the AbstractVE.

At the task level (Figure 4), we see that the overall recall improvement provided by the MixedVE is pertinent for all task types except for the Spatial tasks.

Pairwise comparisons reveal statistically significant differences between the VEs except with the Spatial tasks (Table 1). We see that with *all* Visual or Visuospatial task types (including the MapTask), participants' recall accuracy is higher with the MixedVE than with the AbstractVE, and in *most* of them, they also perform better with the MixedVE than with the RealisticVE. For predominantly Visual tasks, participants' recall accuracy with the RealisticVE and the MixedVE is not statistically significant.

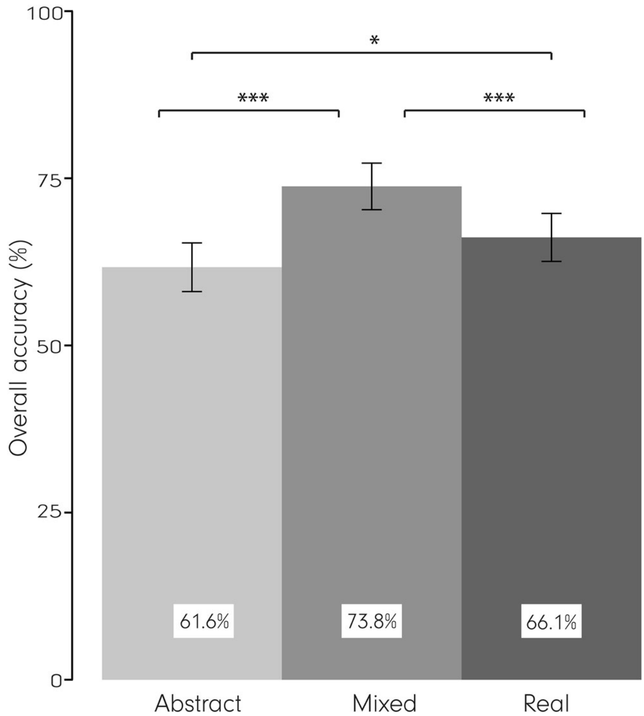


Figure 3. Overall recall accuracy for each VE. Error bars show ±SEM. ***$p < .001$, *$p < 05$.

5.1.1. Individual differences in short-term memory

Based on participants' scores in the MRT (median = 20) and VSM (median = 22) tests, we created high-/low-ability groups using a median split (excluding the median). We call the MRT-based groups *low-MRT* ($n = 19$) and *high-MRT* ($n = 18$), whereas we call the VSM-based groups *low-VSM* ($n = 20$) and *high-VSM* ($n = 19$) from this point forward. Figure 5 shows that there are differences in the recall accuracy of the participants both based on their MRT scores [in favor of the high-MRT with the Realistic VE ($t(36.97) = -2.51$, $p < .05^*$, $r = .38$)] and based on their VSM scores [in favor of the high-VSM for the MixedVE ($t(32.89) = -2.18$, $p < .05^*$, $r = .35$), and the RealisticVE ($t(34.53) = -2.44$, $p < .05^*$, $r = .38$)].

Table 1. Mean recall accuracies, ANOVA (F, p, η_p^2, and pairwise comparisons (for statistically significant results).

Task	Abstract (A) Mean ± SD (%)	Mixed (M) Mean ± SD (%)	Realistic (R) Mean ± SD (%)	Repeated measures ANOVA	Pairwise comparisons
Overall	61.6 ± 12.0	73.8 ± 11.5	66.1 ± 11.9	$F(2,84) = 21.1$, $p < .001$***, $\eta_p^2 = .154$	M–A ($p < .001$)*** M–R ($p < .001$)*** R–A ($p < .05$)*
Visual	56.3 ± 15.8	68.9 ± 15.6	64.3 ± 18.0	$F(2,84) = 9.3$, $p < .001$***, $\eta_p^2 = .092$	M–A ($p < .001$)*** R–A ($p < .05$)*
Visuospatial	63.1 ± 17.4	76.9 ± 17.8	63.8 ± 14.3	$F(2,84) = 16.3$, $p < .001$***, $\eta_p^2 = .129$	M–A ($p < .001$)*** M–R ($p < .001$)***
Map task A (passive)	70.8 ± 23.3	85.1 ± 25.3	72.6 ± 22.0	F(2,84) = 6.7, $p < .01$**, $\eta_p^2 = .069$	M–A ($p < .01$)** M–R ($p < .05$)*
Spatial	72.6 ± 25.2	76.4 ± 21.8	76.4 ± 17.5	$p > .05$	–

Note: We always list the 'winning' VE first (e.g. M–A means the MixedVE led to a higher recall than the AbstractVE). SD: Standard Deviation.

***$p < .001$, **$p < .01$, *$p < .05$, $p < .10$.

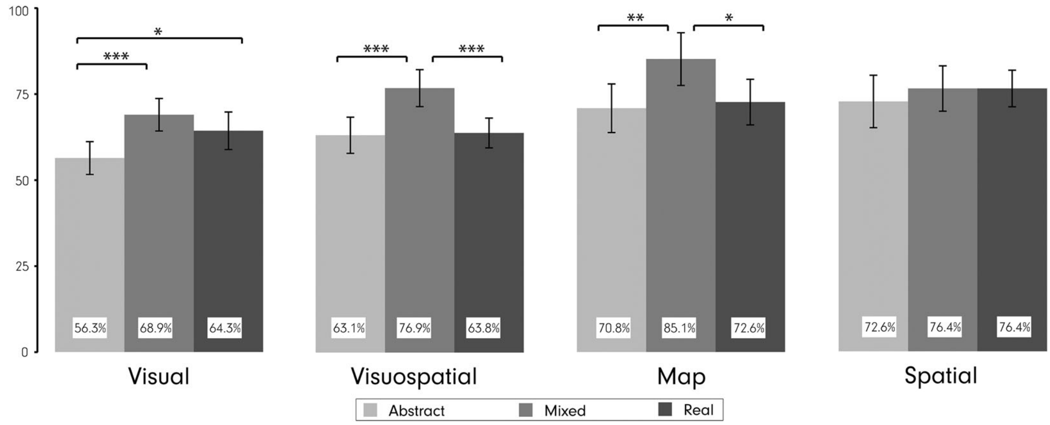


Figure 4. Interactions between visualization types and task types for recall accuracy rates. Error bars show ±SEM. ***$p < .001$, **$p < .01$, *$p < 05$.

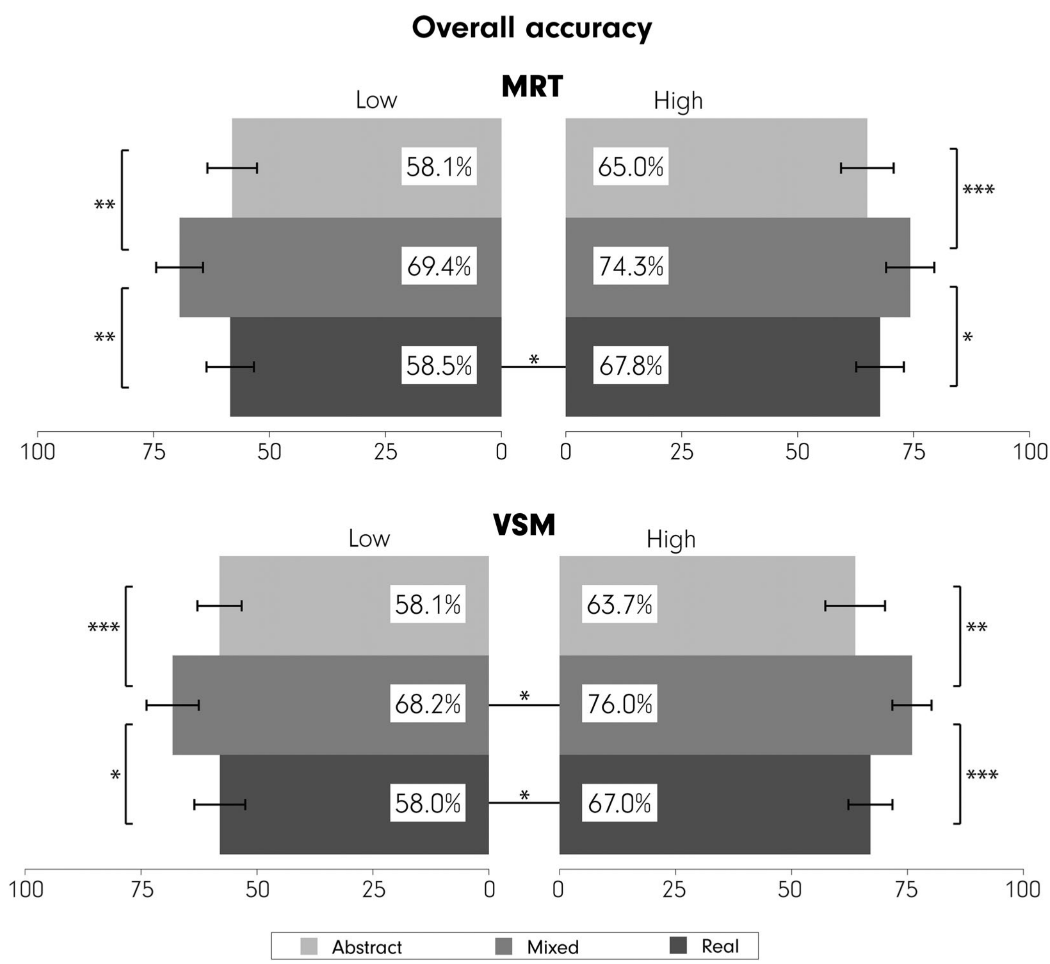


Figure 5. Overall recall accuracy for each visualization type based on MRT- and VSM-split groups. Error bars show ±SEM. ***$p < .001$, **$p < .01$, *$p < 05$.

At the task level, Figure 6 and Table 2 (below) reveal that, for both the MRT and VSM-based groups, irrespective of the abilities, the MixedVE leads to higher recall accuracy than the other two visualization types in *most* tasks. Specifically, we see that the MixedVE offers an advantage over the AbstractVE for both the low and the high-MRT groups for the *Visual tasks*; but not over the RealisticVE. RealisticVE also allows for higher recall accuracy than the AbstractVE for the high-MRT group, but not for the low-MRT group. VSM-split largely confirms these findings for the Visual tasks, except that a high-VSM does not suggest an advantage with the RealisticVE over the AbstractVE. For Visual tasks, the high-VSM group exhibits a higher recall accuracy than the low-VSM group only with the MixedVE ($t(33.52) = -2.38$, $p < .05$*, $r = .38$). For the *visuospatial tasks*, the low-MRT group benefits from the MixedVE more than the Abstract and the RealisticVEs, but for the high-MRT group, visualization type does not make a difference. We also see that the high-MRT group has a higher recall accuracy than the low-MRT group with the Abstract ($t(35.48) = -2.99$, $p < .01$**, $r = .45$) and Realistic ($t(34.27) = -2.68$, $p < .05$*, $r = .42$) VEs, but this difference disappears for the MixedVE. For the same task category, we see that both high- and low-VSM groups benefit from the MixedVE, more than both the Abstract and RealisticVEs. For Visuospatial tasks, we see no differences between the Abstract and RealisticVEs, although the high-VSM group appears to have an advantage with the RealisticVE ($t(34.96) = -2.53$, $p < .05$*, $r = .39$), but not for the Abstract or MixedVEs. For the *MapTask*, MixedVE helps the low-MRT group, and in contrast, the high-VSM group in comparison to the AbstractVE, but we see no differences between the MixedVE and the RealisticVE. However, high-VSM group has a higher recall accuracy than the low-VSM group with the RealisticVE ($t(33.90) = -2.29$, $p < .05$*, $r = .37$). For the *Spatial tasks*, we observe no difference between visualization types in any of the tested conditions.

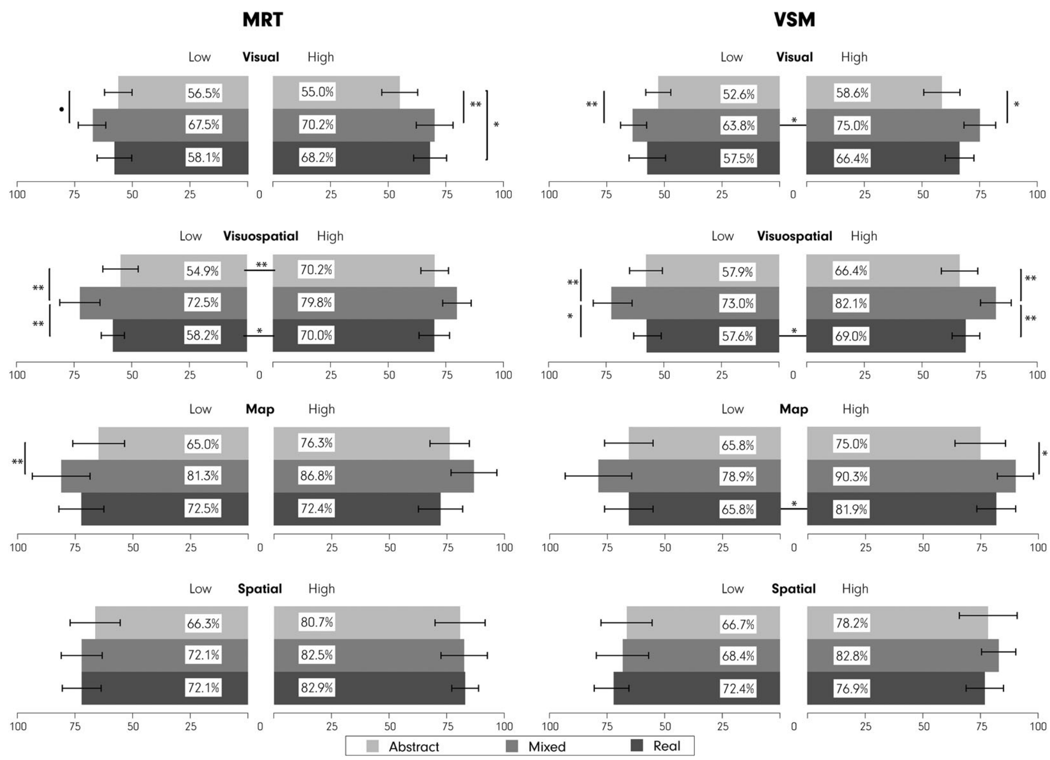


Figure 6. Interactions between the visualization types and task types for the low/high-MRT and low/high-VSM groups' recall accuracy rates. Error bars ±SEM. ***$p < .001$, **$p < .01$, *$p < .05$.

5.2. Long-term memory: comparing recall accuracy in all three stages

Below, we present mean recall accuracies for comparable visualization and task types in all stages (Table 3). We see that the MixedVE continues to facilitate higher recall accuracies than other two VEs in most comparisons *also* in the long-term. The observed differences (MixedVE vs. others) are statistically significant with moderate or large effect sizes. Between the AbstractVE and RealisticVE, we see only one difference at Stage 2, but the p-value only indicates a trend ($p = .0511$), and the effect size is small ($\eta_p^2 = .037$).

In Table 4, we demonstrate the interactions between the task and visualization types over the three stages in terms of *decline* in the recall accuracy. We see an overall decline for all task/visualization types, except that for Visual and MapTasks, AbstractVE does not exhibit a decline in recall performance.

5.2.1. Individual differences

We identified no statistically significant differences ($p > .05$) amongst the three visualization conditions for Stages 2–3, when we group our participants based on their MRT–VSM abilities.

5.3. Visualization preferences

Visualization preferences of the participants *before* and *after* they worked with the VEs are shown in Table 5. We see that before the experiment, majority of the participants preferred the RealisticVE (88%), 12% the MixedVE, whereas *none* preferred the AbstractVE. After the experiment, majority changed their preference to the MixedVE (69%), 31% remained with the RealisticVE, and still none preferred the AbstractVE. Those who changed their preferences all did so from the RealisticVE to MixedVE (there were no instances of the opposite).

To identify whether individual differences changed the preference behavior similarly as in Smallman and Cook's (2011) naive realism studies, we checked the preferences of high/low-MRT and high/low-VSM groups before and after the experiment (Table 6). Unlike in the original naive realism studies, our analyses revealed the same pattern for all groups, irrespective of their spatial abilities or memory capacity.

6. Discussion

Based on previous empirical evidence found in relevant literature, we designed the MixedVE with a texture-highlighting approach, and evaluated it in a three-stage user study. In designing the MixedVE, we made significant adjustments to visual realism levels to lighten cognitive load (Smallman and John 2005; Smallman and Cook 2011), carefully selected the location of the textured buildings to boost memory by placing them at intersections (Röser et al. 2012), and counterbalanced the contents of the textures for saliency (Itti et al. 1998) and memorability (Borkin et al. 2013). In addition to design, individual differences can have an impact in learning performance from visualizations as well as in navigational tasks (e.g. Montello et al. 2004; Huk 2006; Schinazi et al. 2013). Thus, we conducted an analysis of the individual differences in route-recall accuracy based on two measurements: spatial ability (MRT) and visuospatial memory capacity (VSM). Our results provide unique and new insights, and we discuss their implications below.

Our findings overall confirm our main hypothesis that the MixedVE facilitates better route recall than the AbstractVE and RealisticVEs (Figure 3). At the task level (Figure 4), we see that the effectiveness of the MixedVE in route recall is pertinent to all task types except the Spatial tasks. The two tasks we classified 'Spatial' were about orientation (which cardinal direction were you facing at the end of the route), and the number of turns participants' took in the virtual drive. The mean recall accuracy in Spatial tasks is identical for the MixedVE and RealisticVE (76.4%), while it is slightly lower with the AbstractVE (72.6%). These numbers are relatively high in the context of the

Table 2. ANOVAs (F, p, η_p^2 and pairwise comparisons of mean recall accuracies for high/low-MRT and high/low-VSM groups per visualization and task types.

	Tasks	Repeated measures ANOVA	Pairwise comparison	Repeated measures ANOVA	Pairwise comparison
		High-MRT		Low-MRT	
MRT	Overall	$F(2,38) = 10.2$, $p < .001$***, $\eta_p^2 = .101$	M–A ($p < .001$***), M–R ($p < .05$*)	$F(2,40) = 12.0$, $p < .001$***, $\eta_p^2 = .172$	M–A ($p < .01$**) M–R ($p < .01$**)
	Visual	$F(2,38) = 7.0$, $p < .01$**, $\eta_p^2 = .134$	M–A ($p < .01$**), A–R ($p < .05$*)	$F(2,40) = 3.7$, $p < .05$*, $\eta_p^2 = .094$	M–A ($p = .053$·)
	Visuospatial	$F(2,38) = 4.2$, $p < .05$*, $\eta_p^2 = .096$	A–M ($p = .076$·)	$F(2,40) = 10.6$, $p < .001$***, $\eta_p^2 = .172$	M–A ($p < .01$**) M–R ($p < .01$**)
	Map task A	$p > .05$	–	$F(2,40) = 4.1$, $p < .05$*, $\eta_p^2 = .064$	M–A ($p < .01$**)
	Spatial	$p > .05$	–	$p > .05$	–
		High-VSM		Low-VSM	
VSM	Overall	$F(2,36) = 13.7$, $p < .001$***, $\eta_p^2 = .181$	M–A ($p < .01$**), M–R ($p < .001$***)	$F(2,38) = 11.5$, $p < .001$***, $\eta_p^2 = .147$	M–A ($p < .001$***) M–R ($p < .05$*)
	Visual	$F(2,36) = 5.5$, $p < .01$**, $\eta_p^2 = .163$	M–A ($p < .05$*)	$F(2,38) = 3.6$, $p < .05$*, $\eta_p^2 = .091$	M–A ($p < .01$**)
	Visuospatial	$F(2,36) = 9.2$, $p < .001$***, $\eta_p^2 = .174$	M–A ($p < .01$**), M–R ($p < .01$**)	$F(2,38) = 8.8$, $p < .001$***, $\eta_p^2 = .163$	M–A ($p < .01$**) M–R ($p < .05$*)
	Map task A	$F(2,36) = 3.4$, $p < .01$**, $\eta_p^2 = .090$	M–A ($p < .05$*)	$p > .05$	–
	Spatial	$p > .05$	–	$p > .05$	–

Note: We list the 'winning' VE first (e.g. M–A means the MixedVE led to a higher recall than the AbstractVE).
***$p < .001$, **$p < .01$, *$p < .05$, $p < .10$.

Table 3. Mean recall accuracies in all stages, ANOVA (F, p, η_p^2, and pairwise comparisons for statistically significant results.

Task		Abstract (A) Mean ± SD (%)	Mixed (M) Mean ± SD (%)	Realistic (R) Mean ± SD (%)	Repeated measures ANOVA	Pairwise comparisons
Stage 1 (short-term)	Overall	61.6 ± 12.0	73.8 ± 11.5	66.1 ± 11.9	$F(2,84) = 21.1$, $p < .001$***, $\eta_p^2 = .154$	M–A ($p < .001$)*** M–R ($p < .01$)*** R–A ($p < .05$)*
	Visual	56.3 ± 15.8	68.9 ± 15.6	64.3 ± 18.0	$F(2,84) = 9.3$, $p < .001$***, $\eta_p^2 = .092$	M–A ($p < .001$)*** R–A ($p < .05$)*
	Visuospatial	63.1 ± 17.4	76.9 ± 17.8	63.8 ± 14.3	$F(2,84) = 16.3$, $p < .001$***, $\eta_p^2 = .129$	M–A ($p < .001$)*** M–R ($p < .001$)***
	Map task A (passive)	70.8 ± 23.3	85.1 ± 25.3	72.6 ± 22.0	$F(2,84) = 6.7$, $p < .01$**, $\eta_p^2 = .069$	M–A ($p < .01$)** M–R ($p < .05$)*
Stage 2 (long-term 1, 1 h later)	Overall	61.5 ± 21.8	68.8 ± 18.4	57.3 ± 22.6	$F(2,84) = 10.3$, $p < .001$***, $\eta_p^2 = .050$	M–A ($p < .05$*) M–R ($p < .001$***)
	Visual	49.1 ± 24.9	59.5 ± 22.2	52.7 ± 22.5	$p > .05$	–
	Visuospatial	NA	NA	NA	NA	NA
	Map task B (sketching)	67.6 ± 28.6	73.4 ± 27.1	59.7 ± 31.3	$F(2,84) = 8.8$, $p < .001$***, $\eta_p^2 = .037$	M–R ($p < .01$*) A–R ($p = .0511$·)
Stage 3 (long-term 2, one week later)	Overall	54.6 ± 17.0	64.6 ± 18.6	54.5 ± 18.7	$F(2,84) = 21.0$, $p < .001$***, $\eta_p^2 = .064$	M–A ($p < .001$***) M–R ($p < .001$***)
	Visual	56.0 ± 24.4	50.6 ± 18.5	49.1 ± 16.2	$p > .05$	–
	Visuospatial	41.2 ± 15.8	65.8 ± 19.7	47.5 ± 13.7	$F(2,84) = 31.6$, $p < .001$***, $\eta_p^2 = .289$	M–A ($p < .001$)*** M–R ($p < .001$)***
	Map task B (sketching)	64.1 ± 33.0	70.6 ± 33.7	62.4 ± 35.4	$p > .05$	–

Note: Pairwise comparison columns always lists the 'winning' VE first (e.g. M–A means the MixedVE led to a higher recall than the AbstractVE). SD: Standard Deviation. NA: Not available.
***$p < .001$, **$p < .01$, *$p < .05$, $p < .10$.

Table 4. Mean recall accuracies in all stages for comparable tasks in each VE. ANOVA (F, p, η_p^2, and pairwise comparisons.

Visualization type		Stage 1 Mean ± SD (%)	Stage 2 Mean ± SD (%)	Stage 3 Mean ± SD (%)	Repeated measures ANOVA	Pairwise comparison
Visual	Abstract	56.3 ± 15.8	49.1 ± 24.9	56.0 ± 24.4	$p > .05$	–
	Mixed	68.9 ± 15.6	59.5 ± 22.2	50.6 ± 18.5	$F(2,84) = 13.3$, $p < .001$***, $\eta_p^2 = .138$	Stage$_{1-3}$ ($p < .001$***) Stage$_{1-2}$ ($p < .05$*) Stage$_{2-3}$ ($p = .054$·)
	Real	64.3 ± 18.0	52.7 ± 22.5	49.1 ± 16.2	$F(2,84) = 8.9$, $p < .001$***, $\eta_p^2 = .105$	Stage$_{1-2}$ ($p < .05$*) Stage$_{1-3}$ ($p < .001$***)
Visuospatial	Abstract	63.2 ± 17.4	NA	41.2 ± 15.8	–	$t(41) = 6.90$, $p < .001$***, $r = .73$
	Mixed	76.9 ± 17.8	NA	65.7 ± 19.7	–	$t(41) = 2.68$, $p < .05$*, $r = .38$
	Real	63.8 ± 14.3	NA	47.5 ± 13.7	–	$t(41) = 6.36$, $p < .001$***, $r = .70$
Map task	Abstract	70.8 ± 23.4	67.6 ± 28.6	64.1 ± 33.0	$p > .05$	–
	Mixed	85.1 ± 25.3	73.4 ± 27.1	70.6 ± 33.7	$F(2,84) = 4.9$, $p < .01$**, $\eta_p^2 = .046$	Stage$_{1-3}$ ($p < .05$*)
	Real	72.6 ± 22.0	59.7 ± 31.3	62.4 ± 35.4	$F(2,84) = 3.8$, $p < .05$*, $\eta_p^2 = .034$	Stage$_{1-2}$ ($p = .053$·)

Note: Pairwise comparison column lists significant differences between the three stages. 'Winning' stage is listed first. SD: Standard Deviation. NA: Not available.
***$p < .001$, **$p < .01$, *$p < .05$,·$p < .10$.

Table 5. Participants' preferences for the visualization types before and after the experiment.

	Preference before	Preference after	% switched
Abstract	0 (0%)	0 (0%)	–
Mixed	5 (12%)	29 (69%)	0 (0%)
Real	37 (88%)	13 (31%)	24 (65%)

Table 6. High/low-MRT and high/low-VSM groups' preferences for the visualization types before and after the experiment.

	Preference before				Preference after			
	High-VSM		Low-VSM		High-VSM		Low-VSM	
Abstract	0	(0%)	0	(0%)	0	(0%)	0	(0%)
Mixed	1	(5%)	4	(21%)	13	(72%)	14	(74%)
Real	17	(95%)	15	(79%)	5	(28%)	5	(26%)
	High-MRT		Low-MRT		High-MRT		Low-MRT	
Abstract	0	(0%)	0	(0%)	0	(0%)	0	(0%)
Mixed	2	(10%)	3	(15%)	13	(68%)	14	(70%)
Real	17	(90%)	17	(85%)	6	(32%)	6	(30%)

experiment, but not particularly higher or lower than in the other tasks, thus an experimental artifact (such as ceiling or floor effect) does not explain why visualization type did not matter for this task. It might be best explained by the fact that this task essentially requires no visual cues. For the tasks that require the use of visuospatial memory (Visuospatial, MapTask), selectively provided visual cues in the MixedVE improve recall accuracy (by ~10%) compared to both the AbstractVE and RealisticVEs. This pattern is somewhat different for Visual tasks, where we see that the recall accuracy with the RealisticVE competes with the MixedVE, while both VEs with visual cues (Mixed/Realistic) lead to better recall accuracy than the AbstractVE. The fact that the RealisticVE overall facilitates *visual* memory better than the AbstractVE is not surprising, but it is noteworthy that it does not impair the performance in this task type, suggesting that the cognitive load is not 'categorically' too high with fully realistic displays, but it is rather task-specific.

After studying *whether* our proposed MixedVE is effective for route learning (overall recall accuracy shows that it is), and *for what* (analyses at the task level shows it offers benefits mostly in Visuospatial and Visual tasks), we ask *whom* it might benefit most. We expected that participants with high memory capacity (high-VSM) would not be affected as badly from the cognitive load induced by the RealisticVE, especially for the Visual tasks; whereas participants with higher spatial abilities (high-MRT) would do well with tasks with spatial components in them (Spatial, Visuospatial, Map) irrespective of the visualization type. In turn, low-MRT/VSM participants would potentially benefit more from the modifications offered by the MixedVE in all conditions. Overall, our findings show that the MixedVE helps *all* participants (Figure 5), irrespective of their spatial abilities or memory capacities (the RealisticVE and AbstractVEs lead to no differences in performance across MRT/VSM groups). The high-MRT participants overall had a higher recall accuracy than the low-MRT participants with the RealisticVE (9.3% difference). This might mean that high-MRT participants are able to bypass the cognitive overload introduced by the RealisticVE better than the low-MRT, but AbstractVE is also hard for the high-MRT. Memory capacity (VSM-split) did not matter for the recall accuracy with the AbstractVE either, but we see that the high-VSM benefit more than the low-VSM from the MixedVE (by 7.8%) *and* the RealisticVE (by 9%). The VSM (memory capacity) matters clearly for tasks that are of visual/visuospatial nature. Overall, these findings confirm that having a larger capacity for spatial abilities or memory gives participants advantages in some conditions (Wolbers and Hegarty 2010), but the MixedVE improves everyone's route learning performance.

An in-depth analysis of the interactions between individual differences, visualization and task types reveal that, except in Spatial tasks, MixedVE offers benefits in *most* tested conditions against the AbstractVE, and in *some* against the RealisticVE, irrespective of spatial abilities or memory capacity (Figure 6). For the Spatial tasks, varying visual realism seems to be irrelevant *also* irrespective of spatial abilities or memory capacity. For the other tasks (Visual/Visuospatial/Map), most notably, descriptive statistics suggest in *all* cases MixedVE improves performance. Some of the differences are not statistically significant, however, note that we split the participants into groups of $n \cong 20$ based on their MRT/VSM scores (thus the sample size might hinder identifying some differences that are there). Statistically significant differences suggest that MixedVE improves route learning performance for the low-MRT participants in majority of the cases, whereas it helps the high-MRT participants only with the Visual tasks. Reviewing the VSM-based results, we see that MixedVE improves route learning performance more often for the high-VSM participants, but also for the low-VSM participants in two task types. Also interestingly, RealisticVE does not appear to impair performance severely (i.e. not statistically significantly) in many cases when compared to other visualization types, but when we compare the groups of high- vs. low-MRT/VSM, we see that in three cases, the high-ability group outperforms the low-ability group with the RealisticVE (high-MRT in Visuospatial, high-VSM in Visuospatial and MapTasks). In these cases, there are no group differences for the MixedVE, which suggests that the MixedVE brings the performance of the lower-ability participants on par with the higher-ability participants. These findings are consistent with our expectations based on previous work and the results are desirable, given that we often want to create designs that work for all.

Since we were set out to test *learning*, we examined if the MixedVE's benefits would persist over time. It is clear that we gradually forget what we learn (Luck and Hollingworth 2008). Our findings also indicate a steady decline in recall accuracy in Stages 2–3 for the MixedVE and the RealisticVE. The AbstractVE appears to have constant recall levels across all stages for the Visual and Map tasks: for the Visual tasks this is not a surprise, given that the visual cues are important for this task type, and without the visual cues the task is too hard from the beginning (~50% recall accuracy is close to 'chance'). For the Map tasks, the reasons might be more complex: the decline for the AbstractVE is not statistically significant for the Map task (Table 4), possibly because participants predominantly need to perform a perspective switch and the visual cues may not be *as critical*. However, the MixedVE continues to facilitate better recall accuracies than the other VEs also in the long-term (Table 3). Interestingly, the differences between spatial abilities and memory capacity in Stage1 disappear over time; suggesting that higher cognitive abilities help in short-term tasks, but do not necessarily assist in long-term recall of learned routes.

Our analysis of participants' visualization *preferences* (Tables 5–6) shows that the RealisticVE is popular at first, but after working with the VEs, majority prefers the MixedVE. This finding contradicts Smallman and Cook's (2011) observation that (especially the low-spatial) participants do not seem to realize which visualization assists them. Our participant's 'zero interest' in the AbstractVE and initially strong preference toward the RealisticVE supports that realism is generally more attractive, but similarly to some previous work (e.g. Brügger et al. 2016), they are able to detect what assists them once they worked with the visualizations, irrespective of their cognitive abilities.

7. Conclusions

For our proposed 'MixedVE', we adjusted the levels of visual realism, and deliberately selected the location of photo-textures to serve as memorable landmarks. Our rigorous evaluation demonstrates that the design principles we adopted in creating the MixedVE indeed facilitate route learning better than an AbstractVE and a RealisticVE. This observation remained overall true when we scrutinized the possible moderating factors (task types and cognitive abilities). MixedVE consistently led to comparatively higher recall accuracies (and never impaired performance); benefiting all participants irrespective of their cognitive abilities, both in short- and long-term.

Our overall aim is contributing toward empirically verified design guidelines for creating memorable GeoVEs, specifically to assist people to better memorize routes. We believe our findings will be relevant to VR content creators, GIScience and spatial cognition researchers, and has the potential to improve the navigation experience in real world if used as a training device. While in this paper our main interest was in the design and use of the VEs; in future experiments, further group differences (e.g. effects of age) can be examined, and real world navigation performance of individuals can be studied after training them with the MixedVE in comparison to a group trained with the RealisticVE, to confirm MixedVE's utility and usefulness as a 'memory training device'.

Acknowledgements

We would like to thank our participants for their valuable time and our reviewers for their valuable input.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Swiss National Science Foundation (Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung) [grant number 200021_149670] (SNSF project VISDOM).

ORCID

Ismini E. Lokka http://orcid.org/0000-0001-7970-1106
Arzu Çöltekin http://orcid.org/0000-0002-3178-3509

References

Bektaş, Kenan, and Arzu Çöltekin. 2011. "An Approach to Modeling Spatial Perception for Geovisualization." *Procedia – Social and Behavioral Sciences* 21 (January): 53–62. doi:10.1016/j.sbspro.2011.07.027.

Bernabé Poveda, Miguel Angel, and Arzu Çöltekin. 2015. "Prevalence of the Terrain Reversal Effect in Satellite Imagery." *International Journal of Digital Earth*, 8 (8): 640–655. http://www.tandfonline.com/doi/abs/10.1080/17538947.2014.942714?journalCode=tjde20.

Betrancourt, Mireille. 2005. "The Animation and Interactivity Principles in Multimedia Learning." In *The Cambridge Handbook of Multimedia Learning*, edited by R. Mayer, 287–296. Cambridge: Cambridge University Press. https://www.cambridge.org/core/books/cambridge-handbook-of-multimedia-learning/65ED189F0FD10F10A51ADAFF9AE14400.

Biland, Julien, and Arzu Çöltekin. 2017. "An Empirical Assessment of the Impact of the Light Direction on the Relief Inversion Effect in Shaded Relief Maps: NNW Is Better Than NW." *Cartography and Geographic Information Science* 44 (4): 358–372. http://www.tandfonline.com/doi/abs/10.1080/15230406.2016.1185647.

Bleisch, Susanne, and Jason Dykes. 2014. "Quantitative Data Graphics in 3D Desktop-based Virtual Environments – An Evaluation." *International Journal of Digital Earth*, 8 (8): 623–639. http://www.tandfonline.com/doi/abs/10.1080/17538947.2014.927536.

Borkin, Michelle A., Azalea A. Vo, Zoya Bylinskii, Phillip Isola, Shashank Sunkavalli, Aude Oliva, and Hanspeter Pfister. 2013. "What Makes a Visualization Memorable." *IEEE Transactions on Visualization and Computer Graphics* 19: 2306–2315. doi:10.1109/TVCG.2013.234.

Brügger, Annina, Sara Irina Fabrikant, and Arzu Çöltekin. 2016. "An Empirical Evaluation of Three Elevation Change Symbolization Methods Along Routes in Bicycle Maps." *Cartography and Geographic Information Science*, 1–16. doi:10.1080/15230406.2016.1193766.

Brychtová, Alžběta, and Arzu Çöltekin. 2017. "The Effect of Spatial Distance on the Discriminability of Colors in Maps." *Cartography and Geographic Information Science* 44 (3): 229–245. doi:10.1080/15230406.2016.1140074.

Bülthoff, Heinrich H., Jennifer L. Campos, and Tobias Meilinger. 2008. "Virtual Reality as a Valuable Research Tool for Investigating Different Aspects of Spatial Cognition." In *International Conference on Spatial Cognition*, 1–3. Berlin: Springer.

Christou, Chris G., and Heinrich H. Bülthoff. 1999. "View Dependence in Scene Recognition After Active Learning." *Memory & Cognition* 27 (6): 996–1007. doi:10.3758/BF03201230.

Claramunt, Christophe, and Stephan Winter. 2007. "Structural Salience of Elements of the City." *Environment and Planning B: Planning and Design* 34 (6): 1030–1050. doi:10.1068/b32099.

Cockburn, Andy, and Bruce McKenzie. 2002. “Evaluating the Effectiveness of Spatial Memory and 3D Physical and Virtual Environments.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 203–2103. Minneapolis, MN: ACM.

Çöltekin, Arzu, Rebecca Francelet, Kai-Florian Richter, John Thoresen J., and Sara Irina Fabrikant. 2017. “The Effect of Visual Realism, Spatial Abilities, and Competition on Performance in Map-based Route Learning in Men.” *Cartography and Geographic Information Science*. Advance online publication. doi:10.1080/15230406.2017.1344569.

Çöltekin, Arzu, Ismini E. Lokka, and Martin Zahner. 2016. “On the Usability and Usefulness of 3D (Geo)Visualizations – A Focus on Virtual Reality Environments.” In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Commission*, 387–392. Prague: ISPRS. http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLI-B2/387/2016/.

Cowan, Nelson. 2001. “The Magical Number 4 in Short Term Memory: A Reconsideration of Mental Storage Capacity.” *Behavioral and Brain Sciences* 24: 87–114. doi:10.1017/S0140525X01003922.

Cowan, Nelson. 2008. “What Are the Differences Between Long-term, Short-term, and Working Memory?” *Progress in Brain Research* 169: 323–338.

Della Sala, Sergio, Colin Gray, Alan Baddeley, Nadia Allamano, and Lindsey Wilson. 1999. “Pattern Span: A Tool for Unwelding Visuo-spatial Memory.” *Neuropsychologia* 37 (10): 1189–1199. doi:10.1016/S0028-3932(98)00159-6.

Dong, Weihua, and Hua Liao. 2016. “Eye Tracking to Explore the Impacts of Photorealistic 3D Representations in Pedestrian Navigation Performance.” *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 41:641–645.

Dünser, Andreas, Karin Steinbügl, Hannes Kaufmann, and Judith Glück. 2006. “Virtual and Augmented Reality as Spatial Ability Training Tools.” In *Proceedings of the 7th ACM SIGCHI New Zealand Chapter’s International Conference on Computer–Human Interaction: Design Centered HCI*, 125–132. Christchurch: ACM.

Ekstrom, Ruth B., John W. French, Harry H. Harman, and D. Dermen. 1976. *Manual for Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service.

Ellis, Paul D. 2010. *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*. New York: Cambridge University Press.

Fisher, Peter, and David Unwin. 2001. *Virtual Reality in Geography*. London: CRC Press.

Freund, Richard D., John W. Brelsford Jr., and Richard C. Atkinson. 1969. “Recognition vs. Recall: Storage or Retrieval Differences?” *The Quarterly Journal of Experimental Psychology* 21 (3): 214–224.

Garlandini, Simone, and Sara Irina Fabrikant. 2009. “Evaluating the Effectiveness and Efficiency of Visual Variables for Geographic Information Visualization.” In *International Conference on Spatial Information Theory*, 195–211. Berlin: Springer.

Gillner, Sabine, and Hanspeter A. Mallot. 1998. “Navigation and Acquisition of Spatial Knowledge in a Virtual Maze.” *Journal of Cognitive Neuroscience* 10 (4): 445–463. doi:10.1162/089892998562861.

Gore, Al. 1998. “The Digital Earth: Understanding Our Planet in the 21st Century.” *Australian Surveyor* 43 (2): 89–91. doi:10.1080/00050348.1998.10558728.

Hegarty, Mary, Matt S. Canham, and Sara I. Fabrikant. 2010. “Thinking About the Weather: How Display Salience and Knowledge Affect Performance in a Graphic Inference Task.” *Journal of Experimental Psychology. Learning, Memory, and Cognition* 36 (1): 37–53. doi:10.1037/a0017683.

Hegarty, Mary, Harvey S. Smallman, and Andrew T. Stull. 2012. “Choosing and Using Geospatial Displays : Effects of Design on Performance and Metacognition.” *Journal of Experimental Psychology: Applied* 18 (1): 1–17.

Hegarty, Mary, Harvey S. Smallman, Andrew T. Stull, and Matt S. Canham. 2009. “Naïve Cartography: How Intuitions About Display Configuration Can Hurt Performance.” *Cartographica: The International Journal for Geographic Information and Geovisualization* 44 (3): 171–186. doi:10.3138/carto.44.3.171.

Hegarty, Mary, and David Waller. 2004. “A Dissociation Between Mental Rotation and Perspective-taking Spatial Abilities.” *Intelligence* 32 (2): 175–191. doi:10.1016/j.intell.2003.12.001.

Huk, Thomas. 2006. “Who Benefits from Learning with 3D Models? The Case of Spatial Ability.” *Journal of Computer Assisted Learning* 22 (6): 392–404. doi:10.1111/j.1365-2729.2006.00180.x.

Itti, Laurent, Christof Koch, and Ernst Niebur. 1998. “A Model of Saliency-based Visual Attention for Rapid Scene Analysis.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (11): 1254–1259. doi:10.1109/34.730558.

Kattenbeck, Markus. 2015. “Empirically Measuring Salience of Objects for Use in Pedestrian Navigation.” In *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*. Seattle, WA: ACM.

Klauer, Karl Christoph, and Zengmei Zhao. 2004. “Double Dissociations in Visual and Spatial Short-Term Memory.” *Journal of Experimental Psychology: General* 133 (3): 355–381. doi:10.1037/0096-3445.133.3.355.

Klippel, Alexander, and Stephan Winter. 2005. “Structural Salience of Landmarks for Route Directions.” In *International Conference on Spatial Information Theory*, 347–362. Berlin: Springer.

Liao, Hua, Weihua Dong, Chen Peng, and Huiping Liu. 2016. "Exploring Differences of Visual Attention in Pedestrian Navigation When Using 2D Maps and 3D Geo-browsers." *Cartography and Geographic Information Science* 0 (0): 1–17. doi:10.1080/15230406.2016.1174886.

Lobben, Amy K. 2004. "Tasks, Strategies, and Cognitive Processes Associated with Navigational Map Reading: A Review Perspective." *The Professional Geographer* 56 (2): 270–281.

Lokka, Ismini, and Arzu Çöltekin. 2016. "Simulating Navigation with Virtual 3D Geovisualizations – A Focus on Memory Related Factors." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives* 41: 671–673.

Lokka, Ismini E., and Arzu Çöltekin. 2017. "Remembering What We See : Designing Virtual Environments to Improve Visuo-spatial Recall for Navigation Tasks." In *Proceedings of the 27th International Cartographic Conference*, 387–392. Prague: ISPRS. http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLI-B2/387/2016/.

Loomis, Jack M., and James J. Blascovich. 1999. "Immersive Virtual Environment Technology as a Basic Research Tool in Psychology." *Behavior Research Methods, Instruments, & Computers* 31 (4): 557–564. doi:10.3758/BF03200735.

Luck, Steven J., and A. Hollingworth. 2008. *Visual Memory*. New York, NY: Oxford University Press.

Lynch, Kevin. 1960. *The Image of the City*. Vol. 11. Cambridge, MA: MIT Press.

MacEachren, Alan M., Robert Edsall, Daniel Haug, Ryan Baxter, George Otto, Raymon Masters, Sven Fuhrmann, and Liujian Qian. 1999. "Virtual Environments for Geographic Visualization: Potential and Challenges." In *Proceedings of the 1999 Workshop on New Paradigms in Information Visualization and Manipulation in Conjunction with the Eighth ACM Internation Conference on Information and Knowledge Management*, 35–40. Kansas City, MO: ACM.

MacEachren, Alan M., and Menno-Jan Kraak. 2001. "Research Challenges in Geovisualization." *Cartography and Geographic Information Science* 28 (1): 3–12. doi:10.1559/152304001782173970.

Mayer, Richard E., and Roxana Moreno. 2003. "Nine Ways to Reduce Cognitive Load in Multimedia Learning." *Educational Psychologist* 38 (1): 43–52. doi:10.1207/S15326985EP3801_6.

McMahan, A. 2003. "Immersion, Engagement, and Presence." *The Video Game Theory Reader* 67: 86.

Meijer, Frank, Branko L. Geudeke, and Egon L. van den Broek. 2009. "Navigating Through Virtual Environments: Visual Realism Improves Spatial Cognition." *Cyberpsychology & Behavior* 12 (5): 517–521. doi:10.1089/cpb.2009.0053.

Miller, George A. 1956. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* 63 (2): 81–97. doi:10.1037/h0043158.

Montello, Daniel R., David Waller, Mary Hegarty, and Richardson E. Anthony. 2004. "Spatial Memory of Real Environments, Virtual Environments, and Maps." In *Human Spatial Memory: Remembering Where*, edited by G. L. Allen and D. Haun, 251–285. Mahwah, NJ: Lawrence Erlbaum Associates.

Park, Denise C., Gary Lautenschlager, Trey Hedden, Natalie S. Davidson, Anderson D. Smith, and Pamela K. Smith. 2002. "Models of Visuospatial and Verbal Memory Across the Adult Life Span." *Psychology and Aging* 17 (2): 299–320. doi:10.1037/0882-7974.17.2.299.

Parush, Avi, and Dafna Berman. 2004. "Navigation and Orientation in 3D User Interfaces: The Impact of Navigation Aids and Landmarks." *International Journal of Human–Computer Studies* 61 (3): 375–395. doi:10.1016/j.ijhcs.2003.12.018.

Peters, Stefan, Mathias Jahnke, Christian E Murphy, Liqiu Meng, and Alias Abdul-Rahman. 2017. "Cartographic Enrichment of 3D City Models – State of the Art and Research Perspectives." In *Advances in 3D Geoinformation*, 207–230. Cham: Springer International.

Peterson, Lloyd, and Margaret Jean Peterson. 1959. "Short-Term Retention of Individual Verbal Items." *Journal of Experimental Psychology* 58 (3): 193–198. doi:10.1037/h0049234.

Plesa, Malisa Ana, and William Cartwright. 2008. "Evaluating the Effectiveness of Non-Realistic 3D Maps for Navigation with Mobile Devices." In *Map-based Mobile Services*, edited by L. Meng, A. Zipf, and S. Winter, 80–104. Berlin: Springer.

Raubal, Martin, and Stephan Winter. 2002. "Enriching Wayfinding Instructions with Local Landmarks." *Proceedings of the Second International Conference on Geographic Information Science* 2478 (1): 243–259.

Richter, Kai-Florian, and Stephan Winter. 2014. *Landmarks*. Cham: Springer.

Röser, Florian, Kai Hamburger, Antje Krumnack, and Markus Knauff. 2012. "The Structural Salience of Landmarks: Results from an On-line Study and a Virtual Environment Experiment." *Journal of Spatial Science* 57 (1): 37–50. doi:10.1080/14498596.2012.686362.

Roth, Robert E., Arzu Coltekin, Luciene Delazari, Homero Fonseca Filho, Amy Griffin, Andreas Hall, Jari Korpi, et al. 2017. "User Studies in Cartography : Opportunities for Empirical Research on Interactive Maps and Visualizations." *International Journal of Cartography*, 1–29. doi:10.1080/23729333.2017.1288534.

Salthouse, Timothy A. 2006. *Theoretical Perspectives on Cognitive Aging*. Hillsdale, NJ: Psychology Press.

Sanchez, Christopher A., and Russell J. Branaghan. 2009. "The Interaction of Map Resolution and Spatial Abilities on Map Learning." *International Journal of Human Computer Studies* 67 (5): 475–481. doi:10.1016/j.ijhcs.2008.12.003.

Scheiter, Katharina, Peter Gerjets, Thomas Huk, Birgit Imhof, and Yvonne Kammerer. 2009. "The Effects of Realism in Learning with Dynamic Visualizations." *Learning and Instruction* 19 (6): 481–494. doi:10.1016/j.learninstruc.2008.08.001.

Schinazi, Victor R., Daniele Nardi, Nora S. Newcombe, Thomas F. Shipley, and Russell A. Epstein. 2013. "Hippocampal Size Predicts Rapid Learning of a Cognitive Map in Humans." *Hippocampus* 23 (6): 515–528. doi:10.1002/hipo.22111.

Semmo, Amir, Matthias Trapp, Jan Eric Kyprianidis, and Jürgen Döllner. 2012. "Interactive Visualization of Generalized Virtual 3D City Models Using Level-of-abstraction Transitions." *Eurographics Conference on Visualization* 31: 885–894.

Slocum, Terry A., Connie Blok, Bin Jiang, Alexandra Koussoulakou, Daniel R. Montello, Sven Fuhrmann, and Nicholas R. Hedley. 2001. "Cognitive and Usability Issues in Geovisualization." *Cartography and Geographic Information Science* 28 (1): 61–75. doi:10.1559/152304001782173998.

Smallman, Harvey S., and Maia B. Cook. 2011. "Naive Realism: Folk Fallacies in the Design and Use of Visual Displays." *Topics in Cognitive Science* 3 (3): 579–608. doi:10.1111/j.1756-8765.2010.01114.x.

Smallman, Harvey S., and Mark S. John. 2005. "Naïve Realism: Limits of Realism as a Display Principle." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 49 (17): 1564–1568. doi:10.1177/154193120504901714.

Sutherland, Ivan E. 1965. "The Ultimate Display." In *Multimedia: From Wagner to Virtual*, edited by R. Packer and K. Jordan, 252–256. New York, NY: W.W. Norton & Company.

Vandenberg, Steven G., and Allan R. Kuse. 1978. "Mental Rotations, a Group Test of Three-dimensional Spatial Visualization." *Perceptual and Motor Skills* 47 (2): 599–604. doi:10.2466/pms.1978.47.2.599.

Wilkening, Jan, and Sara Irina Fabrikant. 2011. "How Do Decision Time and Realism Affect Map-Based Decision Making?" In *International Conference on Spatial Information Theory*, 1–19. Berlin: Springer.

Wolbers, Thomas, and Mary Hegarty. 2010. "What Determines Our Navigational Abilities?" *Trends in Cognitive Sciences* 14 (3): 138–146.

Publication II

Lokka, I.E., Çöltekin, A., Wiener, J. M., Fabrikant, S. I., & Röcke, C. (2018). Virtual environments as memory training devices in navigational tasks for older adults. *Scientific Reports*, 8(1). https://doi.org/10.1038/s41598-018-29029-x

SCIENTIFIC REPORTS

OPEN

Virtual environments as memory training devices in navigational tasks for older adults

Received: 19 January 2018
Accepted: 25 June 2018
Published online: 17 July 2018

Ismini E. Lokka[1], Arzu Çöltekin[1], Jan Wiener[2], Sara I. Fabrikant[1] & Christina Röcke[3]

Cognitive training approaches using virtual environments (VEs) might counter age-related visuospatial memory decline and associated difficulties in wayfinding. However, the effects of the visual design of a VE in route learning are not fully understood. Therefore, we created a custom-designed VE optimized for route learning, with adjusted levels of realism and highlighted landmark locations (MixedVE). Herein we tested participants' route recall performance in identifying direction of turn at the intersection with this MixedVE against two baseline alternatives (AbstractVE, RealisticVE). An older *vs.* a younger group solved the tasks in two stages (immediate *vs.* delayed recall by one week). Our results demonstrate that the MixedVE facilitates better recall accuracy than the other two VEs for both age groups. Importantly, this pattern persists a week later. Additionally, our older participants were mostly overconfident in their route recall performance, but the MixedVE moderated this potentially detrimental overconfidence. Before the experiment, participants clearly preferred the RealisticVE, whereas after the experiment, most of the younger, and many of the older participants, preferred the MixedVE. Taken together, our findings provide insights into the importance of tailoring visualization design in route learning with VEs. Furthermore, we demonstrate the great potential of the MixedVE and by extension, of similar VEs as memory training devices for route learning, especially for older participants.

Navigation is a key component of human daily life, both when moving between locations in familiar environments, and when reaching new destinations in unfamiliar environments. Especially in unfamiliar environments, navigation can be a difficult task. Because of the age-related decline of some of the perceptual and cognitive abilities that support navigation, this difficulty increases as people age[1,2]. In this paper, we seek to develop a better understanding of difficulties and facilitators in route learning for older adults. Specifically, we examine the potential of virtual environments (VEs) that are custom-designed for route learning in compensating for age-related decline in navigational skills.

There are numerous technology-driven approaches to assist with wayfinding, and many dedicated devices provide real-time navigation instructions such as mobile phone apps or in-car navigation devices[3]. These devices assist people in navigating in the real world, but they are not necessarily optimized for *learning* novel routes or entire environments[4]. In fact, the real-time assistance might contribute to the decline in the ability to independently navigate, as a large portion of the mental effort is externalized to the device and no active engagement from the user is necessary[5]. This argument would be in line with cognitive aging propositions of "use it or lose it"[6]. In this context, we view VEs as candidate visuospatial *memory training devices*. Such a memory training device might benefit everyone, but might be especially meaningful for those who struggle to remember routes when walking in unfamiliar environments, such as is often the case for older adults[7,8]. Training can lead to improvements in the trained domain, such as spatial abilities[9], route learning performance[10–12], as well as other cognitive skills[13–16]. VEs are widely used for navigation-related cognitive training[17], as they provide a safe, controlled substitute for the real world. As such, user errors pose no harm to people while navigating, display design can be personalized, and one can navigate the virtual route as many times as needed without too much physical effort. VEs, however, also possess various limitations. Importantly, most VEs only provide visual stimulation. Other sensory information typically involved in locomotion in the real world, such as vestibular-, proprioceptive-, and efferent-information, are reduced or non-existent in most VEs[18]. The VE setups that stimulate senses other than

[1]University of Zurich, Department of Geography, Zurich, 8057, Switzerland. [2]Bournemouth University, Department of Psychology, Bournemouth, BH12 5BB, UK. [3]University of Zurich, University Research Priority Program (URPP) "Dynamics of Healthy Aging", Zurich, 8050, Switzerland. Correspondence and requests for materials should be addressed to A.Ç. (email: arzu.coltekin@geo.uzh.ch)

vision remain complex to set up, and can be prohibitively expensive[19]. Additionally, linked to the decades-old "experimental control vs. ecological validity" debate[20], one might question whether the learning that occurs in a VE is applicable in the real world, even though a VE has the advantage to simulate the real world to a larger degree than other laboratory-based cognitive training experiments with very abstract stimuli. We believe VEs are promising memory training devices, because the benefits we listed earlier and the advantages it offers over other training methods outweigh their limitations. Besides, previous research presents evidence that learning with tools (maps, VEs) might be applicable in the real world[17]. Importantly, and in contrast to the real world, VEs can be easily and systematically manipulated, which allows addressing research questions that cannot be addressed in real world settings[21]. For example, by highlighting or suppressing certain visual features in a VE, we can investigate which visual information is most relevant for the memorability of a route, thus informing the development of navigational training tools in the future.

A common proposition found in visualization literature is that the more closely the representation of an object resembles its real-world counterpart, the easier it is to relate to it, and people are therefore more likely to remember it[22–24]. In contrast, realism in visuospatial displays impairs performance in certain spatial tasks, including the memorization of a route from 2D maps due to factors such as cognitive load[25–27]. It is also important to note that, unaware of their impaired performance in perceptual tasks, people consider realistic visualizations attractive, and thus might be misguided in their visualization preferences[28]. This concept is coined as 'naïve realism', and it appears to be particularly relevant for people with lower spatial abilities who do not calibrate their preferences even after working with the visualizations that are better for them (people with higher spatial abilities do)[28]. As spatial abilities decline in older age[29], naïve realism might be an important concept to consider when studying how older adults interact with visualizations.

The potential of VEs as memory training devices for older adults in the context of route learning, and the effects of varying the design of VEs (specifically, optimizing the realism levels and landmark locations) on the memorability of the routes, are poorly understood[7]. Here we address this gap as it has important consequences for the development and design of novel interventions to target the highly relevant ability of successful navigation, and thus independent living. In the following sections, we review the key literature regarding memory decline in older adults, particularly in a navigation context; and investigate the potential of VEs as memory training devices from a visualization design perspective.

Navigation in Older Adults: Remembering, Forgetting, and Training

As mentioned previously, it has been well-documented that aging has a negative effect on navigation performance[7]. Especially in unfamiliar environments, older adults experience greater navigational difficulties than younger adults[7,8,30]. Such difficulties can discourage older adults from exploring new environments, and negatively affect their independence and overall quality of life[31]. These age-related navigation difficulties derive from a decline in the relevant visuospatial abilities and memory capacity, both of which vary widely across individuals[32]. Most memory systems, including visuospatial memory that is necessary for navigation, seem to weaken across the lifespan[33]; and this has been documented both in virtual and real world experiments[34–38]. As memory declines, people make more misattribution errors[39,40], that is, an actual experience of an event may be misplaced in time, place or source when retrieved from memory[41,42]. Misattribution errors are common amongst older adults, especially when there are many things to remember[41]. Additionally, it has been shown that older adults overestimate the accuracy of their memories and they are too confident on specific details of their recent experiences[43]. However, the findings on misattribution errors and memory-related overconfidence in older adults might be context dependent. Existing studies are often limited to memorizing lists of words[44], and in certain cases, to presenting pictures and videos, such as videos of crime-scenes[45,46]. When it comes to route learning in unfamiliar environments, it has been shown that older adults tend to confuse the location of landmarks at critical decision points[47]. At this point, however, we know little about the effects of visual design on misattribution errors and memory-related overconfidence. Thus, identifying the optimal design choices for the features of visuospatial training material that facilitates navigational performance in later life seems warranted.

Visuospatial Displays as Training Devices for Route Learning

Learning is the consequence of a complex interplay between sensation, perception, cognition, and experience[48]. In many learning tasks, visuospatial information processing plays a key role. A number of design decisions on how the visuospatial information is represented might affect memory, and consequently, impair or improve route learning[49–51]. We find that a VE optimized for route learning should be balanced for the *amount*, the *quality* and the *position* of the presented information[52]. Quality related considerations are beyond the scope of this paper[52,53]. In this paper, we focus on the amount and the position of the presented information, and we examine their impact on route recall *in combination* (i.e., not independently).

Cognitive load is one of the strongest arguments against visual realism as a display principle[28]. Controlling for the *amount* of information by varying the levels of realism is one way to address cognitive load in route learning in a VE. Depending on the context, one can also highlight landmarks by using symbols (e.g., arrows, letters, colors, outlining the object) or by placing discrete objects at critical locations serving as landmarks, and it has been shown that such approaches increase their memorability[54]. On the other hand, as mentioned earlier, an important argument in favor of realism is that a high degree of realism might make it easier for people to recognize, name and thus relate to the elements (e.g., trees, benches, windows) on a display[22] as they acquire a meaning[24]. 'Nameability' of items might be helpful in memorizing them, for example, people remember nameable colors better than others[55]. It has been proposed that the verbal memory systems help in such cases, because people do not rely *only* on visuospatial memory systems for key executive functions such as the encoding, storage and recall of information (i.e., the dual channel theory)[56]. However, in learning from visualizations, the question of 'how much information is too much/too little?' remains persistent[57,58]. In the case of a VE, one might use photo-textures

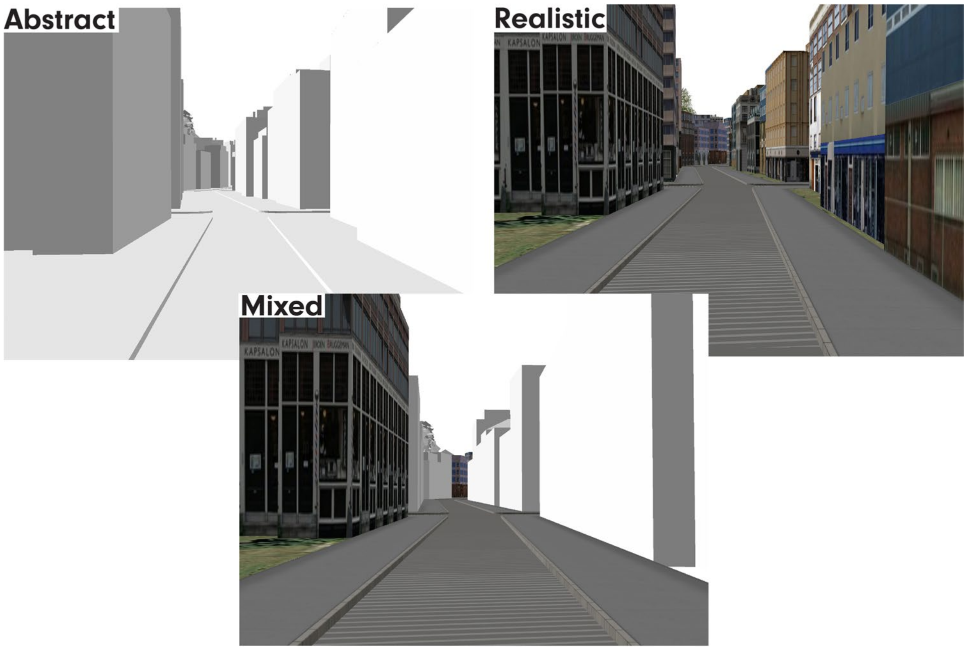


Figure 1. Screenshots from the three VEs to illustrate their visual designs (not to scale). The AbstractVE is rendered in grayscale without photo-textures, whereas the RealisticVE is fully photo-textured. The MixedVE is a combination of the two in which most elements are rendered in grayscale but buildings at critical positions and the road network are photo-textured.

selectively to maintain a sense of realism, and to enable recognition of features, while reducing cognitive load at the same time. On the other hand, a certain level of abstraction guides the attention to task-relevant features[27], which might facilitate remembering and learning.

Besides the amount of information, the *position* of a feature within the scene imposes an important consideration for route learning. Navigation studies and landmark theories mention some unambiguously relevant visuospatial elements that are positioned in specific places in the visual scene for route learning[59–63]: Structural, visual, and/or semantic features determine the importance of landmarks[64–69]. Specifically, *decision points* are critical as in these points people 'take mental notes' of a feature and retain that as a landmark; and reportedly, these features are consistently located in the direction of turn[60]. Related to the position of features, or classes of features, it appears that the structural network (i.e., street network and its spatial layout) provides another important visual anchor in route learning, and might contribute to the memorability of a scene[70].

Our Study

Synthesizing previous work summarized above, we designed an 'optimized VE' for route learning, which we call the MixedVE. The VEs in this study were named based on their relation to abstraction-and-realism; however, note that the optimization is based on *two* important considerations; (a) reducing the *level of realism* by removing photo-textures from task-irrelevant parts of the VE (i.e., manipulating the *quantity* of visual information), and (b) deliberately choosing the locations of the photo-textured elements (i.e., manipulating the *position* of visual information). Because we are interested in *optimizing* the MixedVE as a memory training device, we combined these considerations when designing the MixedVE. Also note that, while we do not investigate aspects of *quality* in this paper, we counterbalance the content of the textures for their semantic qualities based on a previous qualitative assessment for their levels of memorability[52,53].

In sum, in the MixedVE, we highlight selected elements in the scene (i.e., buildings at decision points positioned in the direction of the turn along the route of interest, and the street network) with realistic photo-textures, and suppress the rest by removing photo-textures. Figure 1 shows an illustration of the MixedVE and the other two VEs we used for comparison (AbstractVE and RealisticVE). We chose to compare the MixedVE with a RealisticVE as a high-fidelity representation of the real world. The RealisticVE contains all the visual information including the photo-textures at the navigation-relevant scene elements, however, it does not highlight the navigationally relevant environmental features. The AbstractVE, on the other hand, serves as a baseline condition with no photographic information, and again, no highlighting effect. The fact that the AbstractVE contains considerably less information should significantly reduce the cognitive load induced by photo-textures, although it might increase task difficulty otherwise, because of the lack of anchor points.

We previously demonstrated that younger adults overall benefit from the MixedVE compared to the AbstractVE and RealisticVEs in visual, spatial, and visuospatial memory tasks in a route learning context[53]. In this paper, we examine the potential of the MixedVE as a memory training device in route learning, particularly for older people. Our leading hypothesis is that the MixedVE will successfully serve older people as a memory

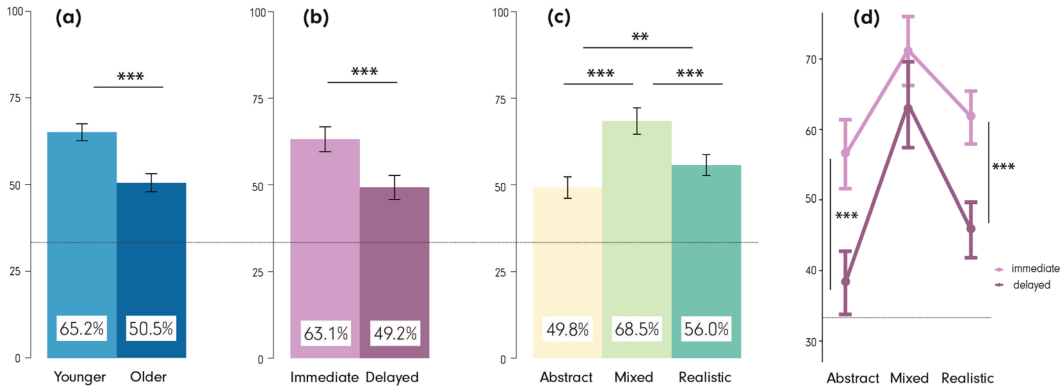


Figure 2. Main effects of (**a**) *age*, (**b**) *recall stage*, and (**c**) *visualization type* on recall accuracy, and (**d**) interactions between *recall stage* × *visualization type* (irrespective of age). The chance level is marked with a light line in 33% recall accuracy. ***p < 0.001, **p < 0.01. Error bars: SEM.

training device in the context of route learning, specifically, in memorizing and identifying the direction of turns at the intersections, because of the following sub-hypotheses:

- Due to the balanced cognitive load and the selective highlighting as a consequence of retaining photo-textures only in navigation-relevant scene elements, irrespective of their age, participants should identify the direction of turn at intersections better (thus, recall the route better) with the MixedVE than with the other VEs, both immediately after the experiment, and a week later.
- Irrespective of age, participants' overall *confidence* in their responses should better align with their recall accuracy with the MixedVE than other VEs. In addition, overall, *older* participants should be *overconfident* in their responses in comparison to younger participants. Thus, the moderating effect of the MixedVE should be more pronounced for the older participant group.
- *Before* the experiment, both older and younger participants should prefer the RealisticVE[25,28]. *After* the experiment, younger participants should change their preferences to the MixedVE. Older participants, however, due to the decline in some of the relevant spatial abilities, might not be able to identify which visualization supports them better, and thus should still prefer the RealisticVE after the experiment.

We tested our hypotheses in a between-subject experiment with an older group (65–75 yrs.) and a younger group (20–30 yrs.) as a comparison group. In the experiment, participants watched a driving simulation video, in which they viewed the route from the 'passenger seat' and were asked to memorize the route. After they watched the videos, participants were given various visuospatial recall tasks in two 'recall stages' (immediate vs. delayed by a week) to measure *learning*. In this paper, we focus on one of the task types; that is, identification of heading direction at intersection points. This is a typical task in route learning studies and previous findings allow us to build our age-related hypotheses for this task type[67,71–74]. In the Procedure section, we describe all of the tasks for full disclosure, and elaborate further on our choice on focusing on this task type. We report the main findings for the two other tasks in the *Appendix: Additional Analysis*. Our three independent variables are *visualization type* (the three VEs), *age* (older vs. younger), and *recall stage* (right after the experiment vs. one week later), whereas we measured three dependent variables: participants' route recall *accuracy*, their *confidence* in their recall performance, and their visualization *preference* before and after the experiment.

Results

We first report the overall route *recall accuracies* of the younger and older participants (age), for the immediate and delayed recall stages (recall stage) with all three VEs (visualization type). Furthermore, we report the *forgetting rates*, (the difference in recall accuracies between the two recall stages) for both groups. We then analyze participants' confidence in their responses. Since confidence in one's success in solving a task can be viewed as one's "perceived accuracy" on that task; we compare the *perceived* and the *actual* accuracies of participants to examine underconfidence or overconfidence (known as *calibration error*[43]). Last but not least, we present participants' visualization *preferences*, and how these preferences shifted among the three VEs before and after the experiment.

Sample size has been estimated via a power analysis using the G-power software. In all tests in which significant results were obtained, the F test was followed by Bonferroni's post-hoc test for multiple comparisons. Associated *p*-values < 0.05 are reported as statistically significant, along with the effect sizes (η_p^2, r, and Cohen's d). Following the convention, we interpret η_p^2 values 0.01, 0.06, 0.14; and Cohen's d values 0.2, 0.5, 0.8; and r values 0.1, 0.3, 0.5 as *small*, *medium* and *large* respectively.

Recall performance based on age, recall stage, and visualization type. A 2 (age) × 2 (recall stage) × 3 (visualization type) mixed-design ANOVA revealed significant differences in recall accuracies for all three independent variables. Figures 2a–c depict the main effects for: (2a) *age*, $F(1, 79) = 29.96$, $p < 0.001$, $\eta_p^2 = 0.10$

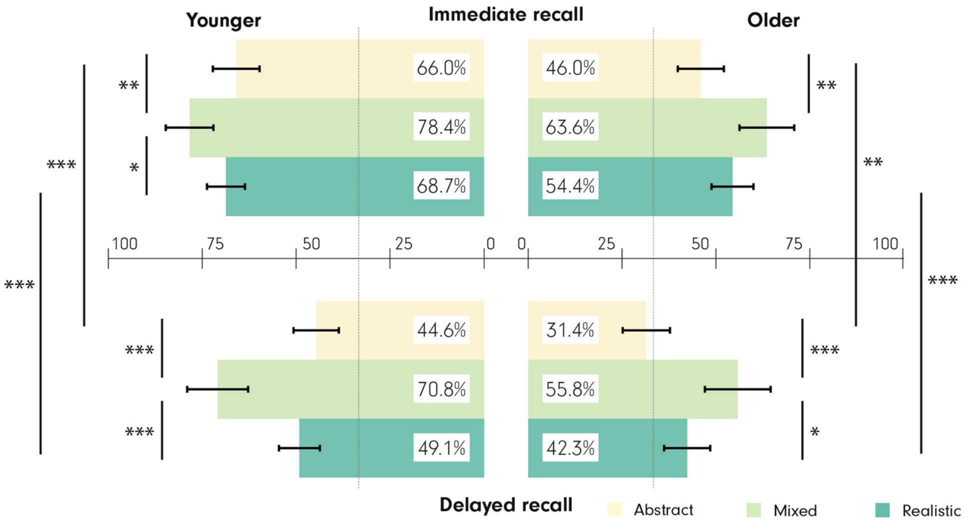


Figure 3. An overview of participants' recall accuracies in the experimental tasks organized by *age*, and *recall stage* for the three visualizations. Left: Younger participants, Right: Older participants. Top: Immediate recall stage, Bottom: Delayed recall stage. The chance level is marked with a light line in 33% recall accuracy. ***p < 0.001, **p < 0.01, *p < 0.05. Error bars: SEM.

(young: 65.2% ± 11.5%, older: 50.5% ± 11.7%); (2b) *recall stage*, $F(1, 79) = 46.17$, $p < 0.001$, $\eta_p^2 = 0.10$ (immediate: 63.1% ± 16.6%, delayed: 49.2% ± 16.0%); and (2c) *visualization type*, $F(2, 158) = 45.78$, $p < 0.001$, $\eta_p^2 = 0.14$ (AbstractVE: 49.8% ± 17.4%, MixedVE: 68.5% ± 19.3%, RealisticVE: 56.0% ± 16.0%). Pairwise comparisons revealed significant differences in participants' recall accuracies between the three VEs. Specifically, recall accuracy was higher with the MixedVE than with the Abstract ($p < 0.001$, $d = 1.02$) and the Realistic VEs ($p < 0.001$, $d = 0.7$), and higher with the RealisticVE than the AbstractVE ($p < 0.01$, $d = 0.37$). Furthermore, the *recall stage* × *visualization type* interaction was significant $F(2, 158) = 3.73$, $p = 0.03$, $\eta_p^2 = 0.01$ (see Fig. 2d). This interaction was driven by greater and statistically significant performance declines between the immediate and delayed recall with the Abstract (18.0%; $t(159.32) = 5.31$, $p < 0.001$, $r = 0.38$) and the Realistic VEs (16.0%, $t(159.93) = 5.42$, $p < 0.001$, $r = 0.39$), while performance decline in the Mixed VE was smaller and non-significant (7.7%; $t(154.26) = 1.89$, $p > 0.05$, $r = 0.15$). None of the other interactions rendered significant results.

Even though we did not observe interactions between *age* × *recall stage* × *visualization type* $F(1, 79) = 0.58$, $p > 0.05$, $\eta_p^2 = 0.00$, we present an overview of the relative recall accuracies of the two age groups in the two stages in Fig. 3. This is accompanied with the inferential statistics in Table 1, to demonstrate how the MixedVE facilitates recall performance better than other VEs in *all* conditions.

Participants' confidence in their recall performance. The calibration error was obtained by dividing the recall accuracies by confidence ratings ("perceived accuracies"). For better readability, we scaled the obtained values to diverge from zero, with zero being the perfect match between perceived and actual recall accuracy, and values diverging in opposite directions from zero signifying overconfidence(o) and underconfidence(u).

A 2 (age) × 2 (recall stage) × 3 (visualization type) mixed-design ANOVA revealed significant differences in participants' calibration errors. The main effects are shown in Fig. 4a–c for (4a) *age* $F(1, 79) = 23.46$, $p < 0.001$, $\eta_p^2 = 0.08$ (younger: 0.03/u ± 0.43, older: 0.21/o ± 0.39), (4b) *recall stage* where there is no significant effect $F(1, 79) = 1.98$, $p > 0.05$, $\eta_p^2 = 0.01$ (immediate: 0.06/o ± 0.32, delayed: 0.11/o ± 0.51), and (4c) *visualization type* $F(2, 158) = 18.17$, $p < 0.001$, $\eta_p^2 = 0.06$ (Abstract: 0.18/o ± 0.41, Mixed: 0.05/u ± 0.48, Realistic: 0.12/o ± 0.36). At the visualization level, pairwise comparisons showed that the calibration errors differed between the MixedVE and the AbstractVE ($p < 0.001$, $d = 0.51$), as well as between the MixedVE and the RealisticVE ($p < 0.001$, $d = 0.4$). Participants are only slightly underconfident with the MixedVE, whereas they are clearly overconfident with the other two VEs. As in recall performance analysis, *recall stage* × *visualization type* interacted: $F(2, 158) = 7.13$, $p < 0.01$, $\eta_p^2 = 0.02$ (Fig. 4d); calibration errors with the MixedVE were close to zero in both stages (immediate: 0 ± 0.31, delayed: 0.09/u ± 0.60), whereas participants exhibited overconfidence both with the AbstractVE (immediate: 0.13/o ± 0.32, delayed: 0.22/o ± 0.49) and the RealisticVE (immediate: 0.04/o ± 0.32, delayed: 0.21/o ± 0.38). Pairwise comparisons show that participants were even *more* overconfident with the RealisticVE in the delayed recall stage ($t(155.32) = 3.14$, $p < 0.01$, $r = 0.24$) than in the immediate recall stage. None of the other interactions rendered significant results.

Similarly as in the recall accuracy analyses, even though the *age* × *recall stage* × *visualization type* 3-way interaction for the calibration error was not statistically significant $F(1, 79) = 1.95$, $p > 0.05$, $\eta_p^2 = 0.00$, we present an exploratory overview of the calibration errors of the two age groups in the two stages in Fig. 5, along with the inferential statistics in Table 2. These results demonstrate that the two age groups may have different calibration

Age	Recall stage	Repeated measures ANOVA	Pairwise comparison
Younger	Immediate	$F(2,84) = 7.80$, $p < 0.001$, $\eta_p^2 = 0.07$	M-A $p < 0.01$**, $d = 0.60$ M-R $p = 0.02$*, $d = 0.51$ R-A $p > 0.05$, $d = 0.15$
	Delayed	$F(2,84) = 20.3$, $p < 0.001$, $\eta_p^2 = 0.22$	M-A $p = 0.001$***, $d = 1.11$ M-R $p < 0.001$***, $d = 0.95$ R-A $p > 0.05$, $d = 0.24$
	Immediate-Delayed (forgetting rate)	(*only pairwise*)	AllVE: $t(41) = 5.36$, $p < 0.001$***, $r = 0.64$ A: $t(41) = 5.49$, $p < 0.001$***, $r = 0.65$ M: $t(41) = 1.33$, $p > 0.05$, $r = 0.20$ R: $t(41) = 5.77$, $p < 0.001$***, $r = 0.67$
Older	Immediate	$F(2,78) = 9.00$, $p < 0.001$, $\eta_p^2 = 0.11$	M-A $p < 0.01$**, $d = 0.82$ M-R $p > 0.05$, $d = 0.44$ R-A $p > 0.05$, $d = 0.45$
	Delayed	$F(2,78) = 13.9$, $p < 0.001$, $\eta_p^2 = 0.16$	M-A $p < 0.001$***, $d = 1.00$ M-R $p = 0.02$*, $d = 0.56$ R-A $p > 0.05$, $d = 0.55$
	Immediate-Delayed (forgetting rate)	(*only pairwise*)	AllVE: $t(38) = 4.10$, $p < 0.001$***, $r = 0.55$ A: $t(38) = 2.89$, $p < 0.01$**, $r = 0.42$ M $t(38) = 1.83$, $p > 0.05$, $r = 0.28$ R: $t(38) = 3.85$, $p < 0.001$***, $r = 0.53$

Table 1. Differences in participants' recall accuracies. In the pairwise comparison column, the VE that facilitates the higher recall accuracy is listed first. M:MixedVE, A:AbstractVE, R:RealisticVE. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

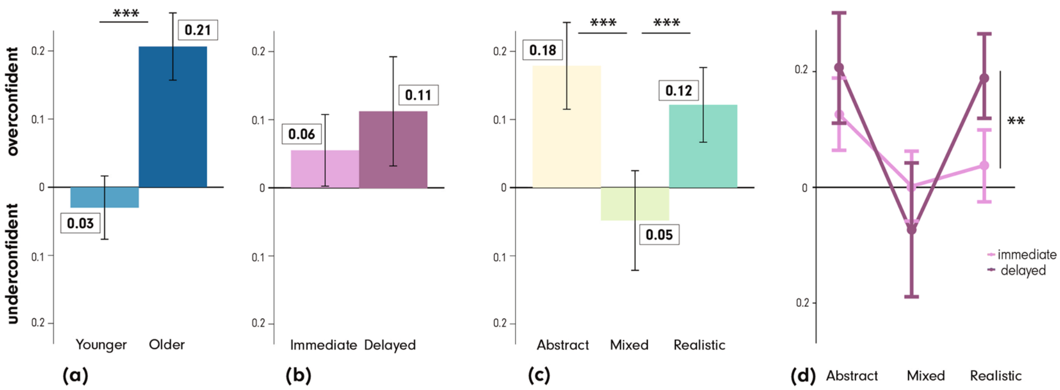


Figure 4. Main effects of (**a**) *age*, (**b**) *recall stage*, and (**c**) *visualization type* on the calibration error, as well as (**d**) interactions between *recall stage* × *visualization type* (irrespective of age). ***$p < 0.001$, **$p < 0.01$. Error bars: SEM.

error patterns. The younger participants rated themselves relatively accurately (they were slightly underconfident) in the immediate stage with all three VEs. In the delayed stage, the younger participants grew overconfident with the Abstract and Realistic VEs, whereas underconfidence persisted with the MixedVE. The older participants were consistently overconfident in all tested conditions, but clearly with the least calibration errors with the MixedVE (close to zero) in both stages. With the lapse of time, both age groups became significantly overconfident with the RealisticVE.

Preference for specific visualization types. Participants' preferences for the three VEs *before* and *after* the experiment are presented in Fig. 6.

As Fig. 6 shows, *before* the experiment, the younger participants mostly preferred the RealisticVE (88%), while only 12% preferred the MixedVE (*none* prefers the AbstractVE). For the older participants, this is even more pronounced: 97% preferred the RealisticVE and the remaining 3% preferred the MixedVE (again none preferred the AbstractVE). *After* the experiment, however, 69% of the younger participants favored the MixedVE, while 31% kept their initial preference for the RealisticVE (the AbstractVE remains unpopular). Older participants display a different pattern: 54% of them still preferred the RealisticVE, a considerable 38% switched to the MixedVE, and 8% preferred the AbstractVE.

The shift in visualization preference from RealisticVE to MixedVE was statistically significant both for the younger ($\chi^2(1) = 67.41$, $p < 0.001$), as well as for the older ($\chi^2(1) = 41.86$, $p < 0.001$) participants. No shift from MixedVE to RealisticVE occurred. The odds ratio (i.e., the effect size) of the younger participants changing their preference from RealisticVE to MixedVE were 16.03 (7.440, 37.090), whereas for the older, this was 22.41 (6.635, 119.180). Due to the unpopularity of the AbstractVE (zero values), we did not include it in the chi-square analysis.

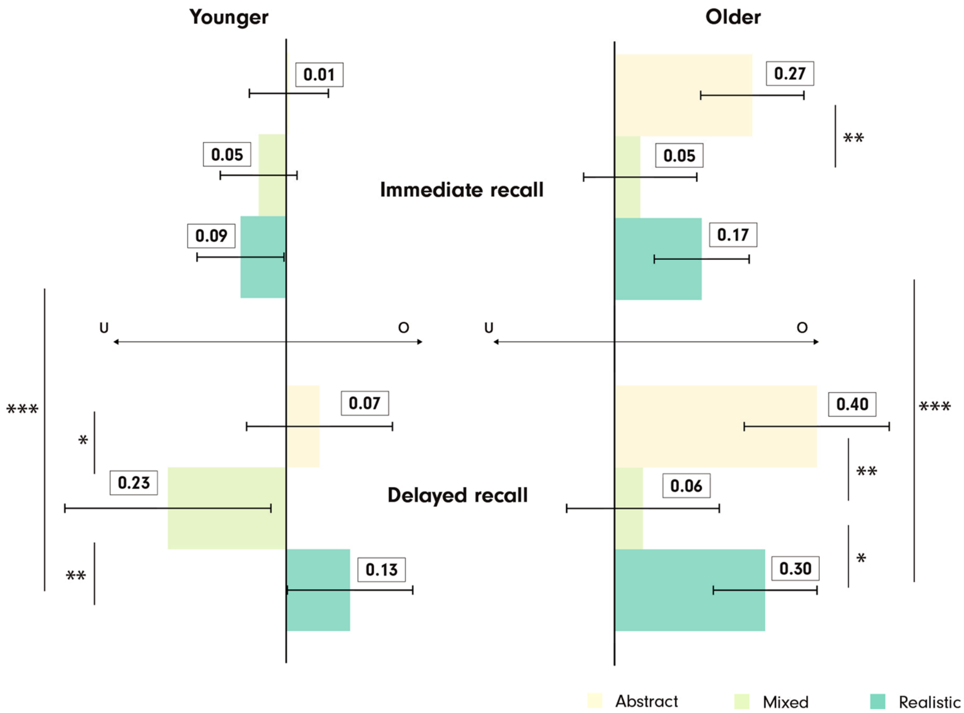


Figure 5. Participants' calibration errors organized by *visualization type*, *age*, and *recall stage*. Left: Younger participants, Right: Older participants. Top: Calibration errors in the immediate recall stage, Bottom: Calibration errors in the delayed recall stage. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. Error bars: SEM. u: underconfidence, o: overconfidence.

Discussion

Despite the popularity and promise of VEs[17], little is known about how older adults are affected by differently-designed VEs in route learning tasks in comparison to younger adults. This is surprising given the importance of maintaining spatial functioning and navigational skills to independently conduct daily life activities across the lifespan. Synthesizing knowledge from several disciplines, we designed an experiment to investigate the potential of a custom-designed MixedVE as a memory training device in route learning. We expected that the MixedVE would help all users; however, because of the age-related decline in visuospatial memory capacity, we were particularly interested in the performance of the older participant group. We included a fully photo-textured RealisticVE as the 'gold standard' because this is a high-fidelity representation of the real world, and an AbstractVE with no photo-textures as a baseline, to examine if, and how much, our customized MixedVE improves the memorability of the given route in comparison to these two VEs.

Route recall accuracy improves with the MixedVE irrespective of age. Our findings clearly confirm that the MixedVE improves recall accuracy of all participants in intersection-by-intersection visuospatial route learning tasks (i.e., in identifying direction of turn) considerably and consistently with large effect sizes; both immediately after the experiment, as well as one week later (Fig. 3). Note that these results are consistent across the 'spatial tasks' as well, whereas we do not observe a clear pattern for the 'visual tasks', possibly because recall accuracies are close to chance level with the visual tasks (see *Appendix: Additional analysis* for overall recall accuracy results based on visual and spatial tasks). Overall, our findings provide clear support for the notion that design decisions are important for successful utilization of VEs for route learning. Note that since we manipulated *two* design elements in the MixedVE — adjusting the level of realism, and deliberately selecting the landmark locations[60] — we cannot distinguish whether and how much each of the two manipulations influence route recall performance. However, the purpose of the study was not to disentangle the contribution of these two design decisions, but to evaluate an "optimized" design. This required us to consider previous knowledge about what design decisions might improve route learning and recall performance. Our findings suggest that reducing the amount of realism while keeping crucial (i.e., navigationally-relevant) information, indeed assists participants in both age groups in identification of the turn directions, and by extrapolation, route recall in general. In other words, since we were set to measure an optimized design against baseline alternatives, we will not discuss the separate effects of realism levels and landmark locations; these were shown by others in dedicated experiments.

Previous work suggests that visualizations that contain too much or too little information can have negative effects on memory performance[28,58]. Our results *regarding Abstract and RealisticVEs* confirm that both too much and too little information indeed impair performance in a VE-based route learning task, and also importantly, this is true also for older adults (65–75 yrs). As mentioned earlier, another key design decision was the position of the highlighted landmarks in a virtual scene. It is known that people rely on landmarks at specific locations in wayfinding tasks[60,70,75]. Our results *with the MixedVE* suggest that 'highlighting' landmarks at task-relevant

Age	Recall stage	Repeated measures ANOVA	Pairwise comparison
Younger	Immediate	$F(2,84) = 1.98, p > 0.05, \eta_p^2 = 0.02$	M-A $p > 0.05, d = 0.23$ M-R $p > 0.05, d = 0.13$ R-A $p > 0.05, d = 0.35$
	Delayed	$F(2,84) = 8.01, p < 0.01, \eta_p^2 = 0.08$	A-M $p = 0.04^*, d = 0.51$ R-M $p < 0.01^{**}, d = 0.65$ A-R $p > 0.05, d = 0.14$
	Immediate-Delayed	(*only pairwise*)	A: $t(41) = 0.76, p > 0.05, r = 0.12$ M: $t(41) = 1.54, p > 0.05, r = 0.23$ R: $t(41) = 3.33, p < 0.01^{**}, r = 0.46$
Older	Immediate	$F(2,78) = 6.94, p < 0.01, \eta_p^2 = 0.07$	M-A $p < 0.01^{**}, d = 0.65$ M-R $p > 0.05, d = 0.37$ R-A $p > 0.05, d = 0.32$
	Delayed	$F(2,78) = 8.63, p < 0.001, \eta_p^2 = 0.10$	M-A $p < 0.01^{**}, d = 0.73$ M-R $p = 0.01^{**}, d = 0.59$ R-A $p > 0.05, d = 0.25$
	Immediate-Delayed	(*only pairwise*)	A: $t(38) = 1.44, p > 0.05, r = 0.23$ M: $t(38) = 0.07, p > 0.05, r = 0.01$ R: $t(38) = 3.00, p < 0.01^{**}, r = 0.44$

Table 2. Differences in participants' calibration errors. In the pairwise comparison column, the VE that leads to the least calibration error is listed first. M:MixedVE, A:AbstractVE, R:RealisticVE. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

locations (i.e., in our case, retaining the photo-textures only in navigation-relevant locations) contributes to intersection-by-intersection route memorization and learning in a VE setup where participants learn passively from a video.

It must be noted that, given that we use VEs as a proxy to real world, using photo-textures for highlighting the features of importance is an appropriate choice, and might transfer well to the real world through the resemblance of detail found in photography. However, the use of photo-textures to highlight the selected features (in this case, buildings and the structural network) is one of the many ways one might design a VE as a memory training device. Other means of highlighting, such as using color or outlining the features of interest, may also prove useful. Therefore, it would be useful to examine other means of highlighting in future experiments for a holistic understanding of highlighting techniques for memory training devices. Also note that the decision to remove realistic detail from a virtual scene immediately triggers the question of *where* such removal would be most appropriate. Removing realistic textures randomly (or based on other criteria) might lead to different outcomes than what we observed in our study. Because we aimed to optimize the MixedVE for route learning, we retained the realistic detail at locations that are relevant to route learning, and for the task examined in this paper, our design decisions provided benefits to the participants.

Our main findings in the recall accuracy analyses confirm an age-related difference disfavoring older adults in route learning performance with medium to large effect sizes[7,8,29,30]. Overall, younger participants recalled routes more accurately than the older participants; irrespective of the visualization type and recall stage (Fig. 3). A closer inspection reveals that *age* and *visualization type* do not interact: Recall performance for both older and younger participants were best with the MixedVE, and the two age groups' recall performance were similar in the two stages. While the age-related memory decline and its various effects on cognitive functions are well documented[47], studies that examine age differences in connection to levels of realism in visuospatial displays are rare. In this study, we observe that the abundance or lack of visual information do not seem to affect the older group *differently* than the younger. Our findings suggest that the complications linked to "too little" and "too much" visual information are fundamental problems that transcend age-related differences.

In contrast to the other two VEs, recall performance did not significantly decline after a week with the MixedVE for either age group (Fig. 3a). This finding is important, because it suggests that removing unnecessary information from a realistic VE *and* leaving it only in navigation-related locations (compare MixedVE vs. RealisticVE); while highlighting relevant information in navigation related locations — in this case, with realistic photo-textures — (compare MixedVE vs. AbstractVE), support *learning* beyond short-term route memorization.

A surprising finding regarding the two recall stages was that the *forgetting rates* of the older participants were not stronger than those of the younger ones after one week. Thus, our findings in the context of route learning in VEs support the notion that age differences in memory are stronger in *encoding* than in *retrieval*, as our older participants did not necessarily experience problems in retrieving the information (stored in their memory) one week later. Evidence regarding age differences in encoding versus retrieval is mixed, however, current understanding is that both are affected by age. An earlier study that tested memory for positions of the pawns in a chess game[76] (a visuospatial task at a different scale) also suggested that it is more the encoding than the retrieval process that is affected by aging. Some other studies, carefully designed to tease apart encoding and retrieval processes experimentally, in contrast, have shown that encoding, retrieval, as well as forgetting rates are negatively affected by aging[77–79]. These differences may depend on a multitude of factors, such as the context in which the studies are conducted or the individual differences among the participants. Further research may help understanding such contradicting observations better.

Overall, as both age groups seem to benefit from the MixedVE, we believe that the basic design assumptions of the MixedVE are fitting choices for route learning in VEs, and that MixedVE, and by extension, similarly designed VEs, have clear potential as memory training devices irrespective of age.

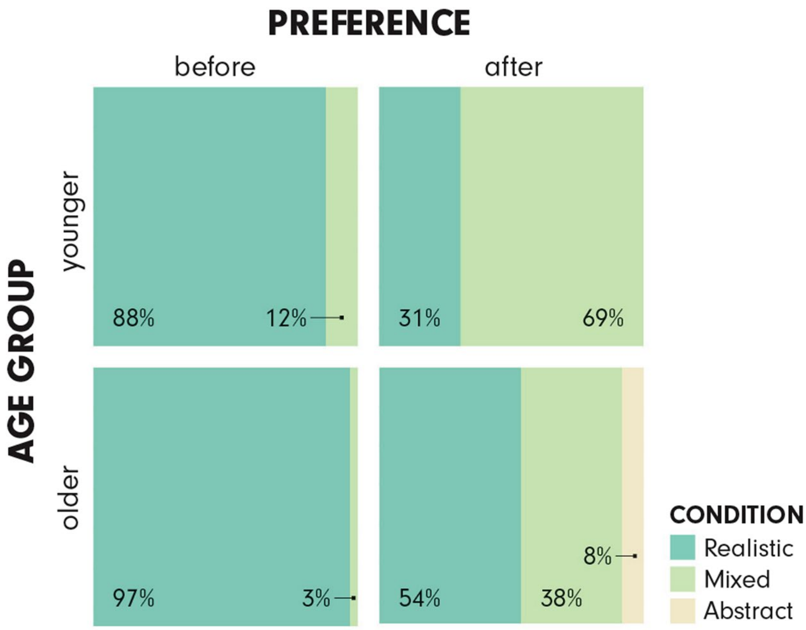


Figure 6. Visualization preferences of younger and older participants before and after the experiment.

Older participants benefit from the MixedVE in calibrating their confidence. It has been previously shown that older people are overconfident in cued recall tasks unrelated to navigation[43]. Thus, we hypothesized that older participants might also be overconfident in route recall tasks. Our calibration error analysis confirms that older participants indeed overestimate their route recall performance in general, in both the immediate and delayed stages with medium effect sizes. This is somewhat alarming, because in a route learning scenario, arguably, overconfidence can be more of a threat than underconfidence. That is, a false belief that one has 'learned the route' might lead to premature action and complications in wayfinding. From this perspective, the fact that the calibration errors with the MixedVE in the delayed recall stage are near-zero for the older group is a very promising result. In other words, with the MixedVE, older people might be less prone to overestimate their performance, and take fewer risks. The younger group is somewhat underconfident with the MixedVE, however, we believe this is less of a threat; as a consequence, they might behave more carefully while navigating after learning with the MixedVE, or practice more.

Both age groups, but particularly the older group seems to be overconfident with the AbstractVE and the RealisticVE in the delayed recall stage with medium effect sizes. With the AbstractVE, the overconfidence may at first appear surprising, as with such low accuracy, one would expect the confidence ratings to be low. Perhaps the visual similarity of the objects to one another, such as it is the case with buildings in the AbstractVE, led to misattribution errors, resulting in a false sense of familiarity a week later when recalling is harder than immediately after the experiment. Similarly, we observe that both age groups had a false belief that they are doing better with the RealisticVE in the delayed recall. This might be explained by the previously documented mismatch in people's accuracy and confidence in other contexts[80]. In this case, because people could identify particular elements in the visual scene after a week passed, they falsely believed that these assisted them to recall a route. Note that the results regarding the calibration error analysis should be viewed as an exploratory analysis, as the overall interaction of *age* × *visualization* × *recall stage* did not reach significance; while these results allow us to hypothesize, more testing is needed to confirm them.

Overall, the MixedVE afforded a better self-assessment than the other two VEs with medium effect sizes for both age groups, possibly because participants could more precisely recall what they have seen. Importantly, the MixedVE offered a clear advantage for the older group, enabling them to calibrate their confidence that matches their performance much better; thus lending itself as a promising candidate for the development of novel training paradigms for all, but especially for older adults.

Participants prefer the RealisticVE before the experiment, but many switch to MixedVE after. We find clear signs of naïve realism[28,81] when participants stated their preferences for the visualization types *before* the experiment. Both age groups overwhelmingly preferred the RealisticVE before the experiment (younger: 88%, older: 97%). These results provide unambiguous evidence of how strongly people are attracted to realistic displays[81].

After participants experienced the VEs and solved the route recall tasks, however, we saw dramatic changes in participants' preferences. As predicted, most of the younger participants shifted their preference from the RealisticVE to the MixedVE after the experiment. This suggests that the younger participants successfully identified the assistance they received from the MixedVE, and valued their performance with it (i.e., sometimes people prefer the inferior product knowingly, simply because they like it). Nonetheless, a notable sub-group of younger participants (31%) stayed with their original preference for the RealisticVE. The older participants' preferences after their experience with the VEs show a different pattern than the younger participants': Even though a large number of older participants also switched to MixedVE (38%), the RealisticVE remained their favorite choice also after the experiment (54%). This may be linked to the overall lower exposure and experience with VE technologies. Furthermore, in Smallman and John's 2011[81] naïve realism study, participants with lower spatial abilities did not necessarily change their preference towards less realistic displays after the experiment, even though those with higher spatial abilities did. Perhaps our findings and theirs are linked; one can speculate that people who do not perform too well for various reasons (age or lower spatial abilities) might be less deliberate about the tools they choose. Thus, when designing future visuospatial memory training devices intended for people with limited experience and abilities, it is important to remember that the acceptance of the proposed device might be a barrier to achieving the memory improvement goals, and additional considerations might be necessary.

Conclusions

Motivated by earlier work on cognitive training, and informed by the principles of visualization design, we tested if one can customize a VE, which could eventually be used as a memory training device in a route learning context. Importantly, because visuospatial memory is negatively affected by age, we focused our efforts on understanding how well our candidate memory training device (the MixedVE) would work for older adults. Specifically, we focused on the visual design of the VE, because design choices can have a strong impact on how well a visualization functions, including its memorability. Thus, we examined aspects of design that should be considered for creating memorable VEs, especially for route learning. Our intuition, as well as the previous work suggested that we represent the world with high fidelity, and replicate the reality in a simulated environment. However, previous empirical evidence in various other contexts led us to believe that we can improve the design of the VE to better function as a memory training device for route learning if we control the amount of visual realism instead. However, we did not 'randomly' remove redundant information. Instead, we designed the MixedVE, in which we used photo-textures only at the navigation-relevant locations, that is, where we knew people would look for landmarks. By 'translating' the previous empirical evidence into design from two perspectives (realism and landmark use), we essentially highlighted navigation-relevant information in the locations that matter to the viewer to increase their saliency and memorability, and we suppressed less relevant information to reduce cognitive load.

Our results provide new insights for the design of VEs and their possible use as visuospatial cognitive training devices for route learning, especially in older adults. Overall, the MixedVE *was* more memorable than the others, and it facilitated high recall accuracy in identification of turn-of-direction tasks at the intersections (and by extension, in route learning), irrespective of age, both in short and long term. The fact that the MixedVE facilitated both immediate and delayed recall and in both age groups shows how effectively the design choices can improve performance whether one is old or young. Furthermore, the stable recall performance with the MixedVE even a week after the participants watched the simulated video (only once), clearly demonstrated its promise as a potential training device. Participants' confidence in their performance matched their actual performance better with the MixedVE compared to the other VEs, and this is especially evident for older participants. The fact that the MixedVE helps with adjusting for overconfidence in older adults has important positive implications on their potential navigational behavior. Furthermore, a large number of participants preferred the MixedVE to others after working with it, even though some more design adjustments might be necessary for an older audience.

Taken together, our findings demonstrate the potential of the MixedVE as a memory training device, for all ages but especially for the older adults, which encourages us to continue this line of research. Aside from these applied implications, we developed a better understanding of the age differences in learning from a VE. Specifically, we know more about the effects of combined visualization design choices (realism levels with landmark locations) on the recall accuracy, confidence and visualization preferences of people from two distinctly different age groups in route recall tasks.

Methods

We conducted a controlled experiment with a mixed factorial ($2 \times 2 \times 3$) design. *Age* was a between-subject factor (younger vs. older), *visualization type* (i.e., Abstract, Mixed, Realistic VEs), and *recall stage* (i.e., immediate vs. delayed) were within-subject factors. All participants performed route learning tasks in all three VEs and at two stages one week apart. As dependent variables, we measured the *recall accuracy* in all the tasks, with a focus on the direction of turns at intersection points, where we also measured participants' *confidence* in their responses, and their visualization *preferences*.

Participants. In total, 81 participants took part in the study: 42 in the younger group (27 ± 2 yrs., 23 female), and 39 in the older group (70 ± 4 yrs., 17 female). The younger participants were between 20–30 years of age and were recruited by word of mouth. The older participants were between 65–75 years of age and were recruited using the participant pool of UZH's University Research Priority Program "Dynamics of Healthy Aging" (http://www.dynage.uzh.ch/en.html). This experiment was approved by the Ethical Committee of the Philosophical Faculty - University of Zurich with the form "Checkliste für die Selbstbeurteilung von Studien auf ethische Unbedenklichkeit". All methods were performed in accordance with the relevant guidelines and regulations and all participants received informed consent, which after agreement they signed. All participants volunteered to participate, signed a written consent form and could withdraw their participation at any time. All participants performed the Mini-Mental State Examination to measure their cognitive status (MMSE). They were included in the study only if they scored a minimum of 27 out of 30[82].

Materials

Apparatus. The experiment was performed in the 3D visualization/virtual reality lab of the GIVA unit of the Department of Geography, of the University of Zurich. Passive drive-throughs of the routes were presented as videos to participants on a large projection screen (230 × 140 cm). The participants were seated at a distance of 2.2 m from the screen to ensure that they could see the whole scene.

Stimuli. Participants were shown videos of drive-throughs in a virtual fictitious city. Using procedural modeling, we designed the city to look as homogenous as possible to control for salient elements which might potentially interfere with route learning. Thus, the city contained buildings and other structures similar in size and architectural style, similar street network (intersection points with ~90 degree angles) over the whole city, and other visual elements (e.g., trees) were also kept similar to each other in size and other visual characteristics. We manipulated the design to obtain three different virtual environments (VEs, visualizations), as illustrated in Fig. 1. These three VEs differed in their degree of realism, and they represent the three main experimental conditions:

- a plain grayscale VE without any photo-textures (AbstractVE)
- a color photo-textured VE (RealisticVE)
- a mix of the two above, in which the buildings at all decision points towards the direction of the turn, and the structural network (street floors) are textured using color photography (MixedVE). Thus, we used photo-textures as a particular *type* of highlighting choice, because we work with realism as an important concept in route learning for transferability of acquired knowledge to the real world; and the *position* of the highlighted landmarks were selected based on landmark theories (we selected the positions that were previously shown as important positions where people took mental notes).

We created two routes in each of these three VEs. Each route consisted of seven intersections (three left, three right turns and one straight). The videos of the drive-through of these two routes were recorded at the same eye-level (1.50 m), had the same duration (100 sec) and were played back at the same speed (30 km/h). Each participant experienced two videos in all three VEs, adding up to a total of six different videos. Videos of the routes were shown only once. Using a Latin squares approach; we systematically rotated the order of the videos.

Task. Participants were instructed to memorize the routes to the best of their ability. Once they watched the video, participants performed a series of different tasks as follows:

(1) Identifying if a scene (screenshot) was on their route ("yes" or "no" answer), based on six screenshots from each path in each VE type (three were correct and three false). We call this set "visual tasks" as the participants would predominantly rely on visual information, while the location of the information was not relevant.
(2) Drawing a sketch of the route using top-down screenshots of each VE. We called this set of tasks "spatial tasks", because location and orientation are the key to solving the task, while the visual information is not as important.
(3) Identifying the direction of turn at each of the seven intersection points based on a screenshot of the intersection point. We called this set of tasks "visuospatial tasks" one has to make use of visual cues, as well as location and orientation to solve the task. In other words, based on previous work[67,72–74,83], we believed that participants would have to rely on both visual and spatial memory to solve this task. We thus considered this predominantly a visuospatial memory task.

More specifically, because participants responded to questions based on three VEs, for two routes in each VE, with a total of seven intersections at each video, they provided a total of 42 individual responses (3 × 2 × 7) in this task set. As mentioned earlier, the intersection points were presented as screenshots from the videos in a randomized order in the recall phase. Participants were asked to choose the direction in which they continued their route among the given options. Approximately one week after the first session (immediate recall stage), participants repeated the tasks without watching the videos again (delayed recall stage). Besides the "I don't know" option, participants could mark *left*, *straight*, and *right*; giving them a 33% chance to guess the correct answer.

Procedure

Upon arrival at the lab, participants signed an informed consent form. We briefed them about the procedure, introduced them to the hardware setup, and answered their questions, if they had any. We then assigned each participant to one of six videos (3 visualization types, 2 routes each). Before starting the actual route learning experiment, we showed participants a representative screenshot from each of the three VEs, and asked them to rate their preferences for a hypothetical route learning task. Immediately after this, the main experiment begun. Participants were given a scenario in which someone took them to a market in an unfamiliar neighborhood, and they were told to memorize the route as they would have to navigate the same route later by themselves. After watching each video only once, they answered a set of recall questions based on this specific video, and rated their confidence for each of their responses using a 5-point Likert scale that varied from "Not at all confident (1)" to "Very confident (5)". After solving the tasks with three of the videos, participants could take a short break for approximately three minutes to counter potential fatigue. The last three videos followed in the same fashion. We then asked the participants which of the three VEs they preferred. There were no time limits in the experiment, thus the experimental duration of the first session (immediate recall stage) varied from 1 h to 1h40min. Participants came back six to eight days after the first session for the second session (delayed recall stage). In this stage, participants were not shown the videos again, thus they responded to the questions based on what they could recall from the first session. The duration of the second session varied from 40min to 1 h.

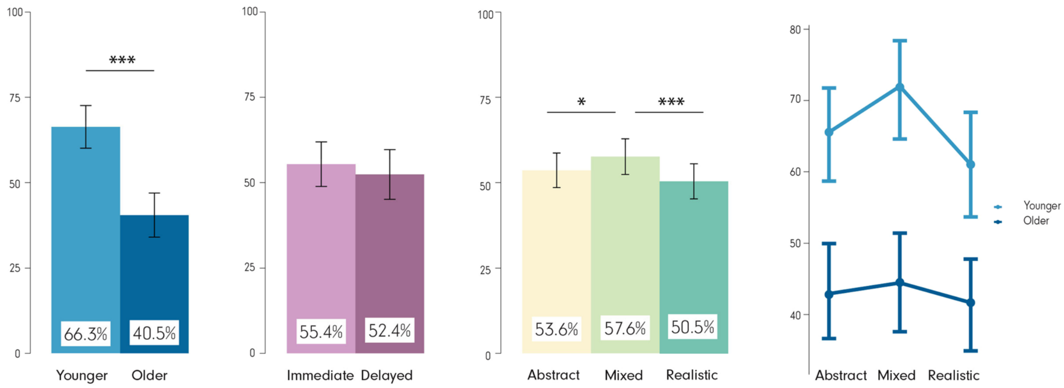


Figure 7. Spatial tasks. Main effects of (**a**) age, (**b**) recall stage, (**c**) visualization type on sketch task, and (**d**) interactions between age × visualization type (irrespective of recall stage). ***p < 0.001, *p < 0.05. Error bars: SEM.

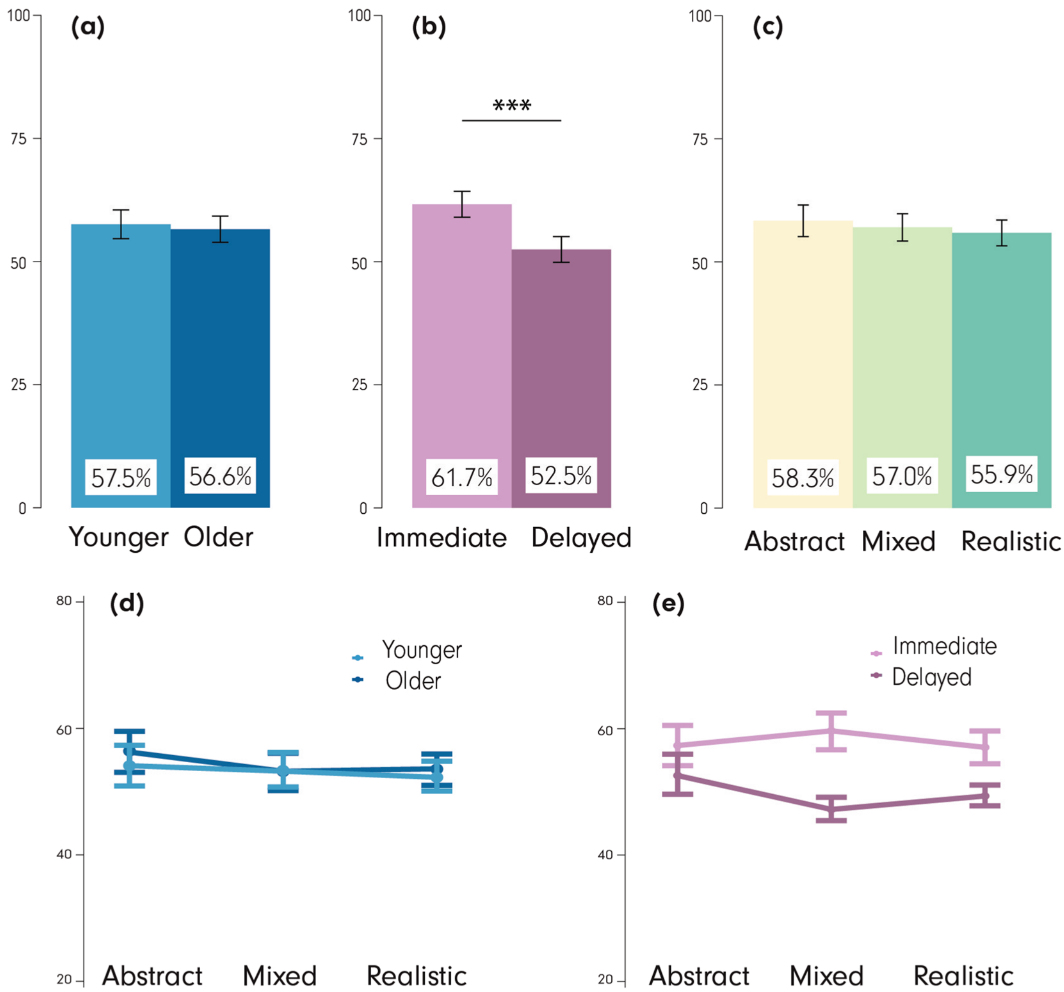


Figure 8. Visual tasks. Main effects of (**a**) age, (**b**) recall stage, (**c**) visualization type on visual task, and interactions between (**d**) age × visualization type (irrespective of recall stage), and (**e**) recall stage × visualization type (irrespective of age). ***p < 0.001. Error bars: SEM.

Appendix: Additional analysis

Spatial task. A 2 (age) × 2 (recall stage) × 3 (visualization) mixed-design ANOVA revealed significant differences in the sketching task for two out of the three independent variables (no difference for recall stage). Figure 7 depicts the descriptive and inferential statistics; statistically significant differences were observed for (a) age $F(1, 79) = 17.04$, $p < 0.001$, $\eta_p^2 = 0.15$ (young: 66.3% ± 31.7%, older: 40.5% ± 30.0%), (b) visualization $F(2, 158) = 11.69$, $p < 0.001$, $\eta_p^2 = 0.01$ (Abstract: 53.6% ± 32.9%, Mixed: 57.6% ± 33.9%, Realistic: 50.5% ± 33.4%) and (c) age × visualization $F(2, 158) = 3.80$, $p < 0.05$, $\eta_p^2 = 0.01$. This interaction was driven by the significantly larger difference in the sketching performance between the Mixed and the Realistic visualizations for the younger participants compared to that of the older (young: 11.0% ± 21.3%, older: 3.0% ± 15.0%, $t(149.35) = 2.77$, $p < 0.01$, $r = 0.22$). Interestingly, the *recall stage* did not reveal statistically significant differences (immediate: 55.4% ± 32.1%, delayed: 52.4% ± 34.7%), neither did any other of the interactions.

Overall, the results from this task are in line with the visuospatial task. Age and visualization seem to matter for the performance, with the MixedVE resulting in best performance compared to both the Abstract and the Realistic VEs. Participants' performance in the different recall stages was not significantly different in the spatial task. This might be explained by the active involvement required to fulfill the task. That is, the fact that participants actively *drew* the path immediately after the experiment, may have resulted in them learning to solve this task better than the other tasks in which they only passively watched the stimuli.

Visual task. A 2 (age) × 2 (recall stage) × 3 (visualization) mixed-design ANOVA revealed significant differences in the visual task for only one of the three independent variables (no differences in age and visualization type, Fig. 8). Statistically significant differences were observed for *recall stage*, $F(1, 79) = 39.71$, $p < 0.001$, $\eta_p^2 = 0.06$ (immediate: 61.7% ± 17.0%, delayed: 52.5% ± 19.1%). Of the interactions, *age × visualization*, ($F(2, 158) = 3.80$, $p < 0.05$, $\eta_p^2 = 0.01$) and *recall stage × visualization interaction*, ($F(2, 158) = 10.05$, $p < 0.001$, $\eta_p^2 = 0.03$) rendered significant results. The *age × visualization* interaction was driven by the significantly larger difference in visual recall performance between the Mixed with the Abstract visualization for the younger participants compared to that of the older (young: 3.6% ± 26.8%, older: −6.7% ± 23.7%, $t(159.65) = 2.59$, $p < 0.05$, $r = 0.20$ [note the (−) in the older recall performance signifies higher recall for the Abstract compared to the MixedVE]). The *recall × visualization* interaction was driven by the significantly larger differences in the visual task performance between the Mixed and the Abstract visualization in the immediate compared to that of the delayed recall (immediate: 5.8% ± 20.7%, delayed: −8.5% ± 28.4%, $t(146.32) = 3.66$, $p < 0.001$, $r = 0.29$), and the Abstract with the Realistic visualization (immediate: −3.2% ± 19.7%, delayed: 8.2% ± 26.2%, $t(148.57) = 3.13$, $p < 0.01$, $r = 0.25$). Interestingly, *age* (young: 57.5% ± 19.5%, older: 56.6% ± 17.7%) and *visualization* (Abstract: 58.3% ± 20.8%, Mixed: 57.0% ± 18.0%, Realistic: 55.9% ± 16.9%) did not reveal statistically significant differences.

Note that regarding the age differences, the literature suggests that spatial memory functioning tends to decline with age, but visual memory might be 'spared'[84]. Therefore, we have speculated from the start that the results for the visual task would be different than other memory tasks, which load on spatial memory more heavily. Initially, the results from the visual task do not seem to agree with the results from the visuospatial and the spatial tasks. Especially the interaction between *age × visualization* shows a conflicting pattern for the two age groups, with the MixedVE being more supportive for the young but not for the older, who seem to achieve higher recall with the AbstractVE. When examining the exact performance values from the visual task, however, we see that the performance is close to the chance level (50%) for the older participants. Thus, lack of interactions in this case may be due to task difficulty, which likely caused a "floor effect" for the older group and for both age groups in the delayed recall stage. In other words, overall task difficulty could have overshadowed these interactions.

Data availability. The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Wolbers, T., Dudchenko, P. A. & Wood, E. R. Spatial memory-a unique window into healthy and pathological aging. *Front. Aging Neurosci.* **6**, 35 (2014).
2. Lester, A. W., Moffat, S. D., Wiener, J. M., Barnes, C. A. & Wolbers, T. The Aging Navigational System. *Neuron* **95**, 1019–1035 (2017).
3. Klippel, A. *et al.* Direction concepts in wayfinding assistance systems. *Work. Artif. Intell. Mob. Syst.*, 1–8 (2004).
4. Münzer, S., Zimmer, H. D. & Baus, J. Navigation assistance: A trade-off between wayfinding support and configural learning support. *J. Exp. Psychol. Appl.* **18**, 18–37 (2012).
5. Parush, A. & Berman, D. Navigation and Orientation in 3D user Interfaces: The Impact of Navigation Aids and Landmarks. *Int. J. Hum. Comput. Stud.* **61**, 375–395 (2004).
6. Hultsch, D. F., Hertzog, C., Small, B. J. & Dixon, R. A. Use it or lose it: Engaged lifestyle as a buffer of cognitive decline in aging? *Psychol. Aging* **14**, 245–263 (1999).
7. Moffat, S., Zonderman, A. B. & Resnick, S. M. Age differences in spatial memory in a virtual environment navigation task. *Neurobiol. Aging* **22**, 787–796 (2001).
8. Muffato, V., Della Giustina, M., Meneghetti, C. & De Beni, R. Age-related differences in pointing accuracy in familiar and unfamiliar environments. *Cogn. Process.* **16**, 313–317 (2015).
9. Uttal, D. H. *et al.* The malleability of spatial skills: A meta-analysis of training studies. *Psychol. Bull.* **139**, 352–402 (2013).
10. Mitolo, M., Borella, E., Meneghetti, C., Carbone, E. & Pazzaglia, F. How to enhance route learning and visuo-spatial working memory in aging: a training for residential care home residents. *Aging Ment. Health* **21**, 562–570 (2017).
11. Lövdén, M. *et al.* Spatial navigation training protects the hippocampus against age-related changes during early and late adulthood. *Neurobiol. Aging* **33**, 620–e9 (2012).
12. Hötting, K., Holzschneider, K., Stenzel, A., Wolbers, T. & Röder, B. Effects of a cognitive training on spatial learning and associated functional brain activations. *BMC Neurosci.* **14**, 73 (2013).
13. Martin, M., Clare, L., Altgassen, A. M., Cameron, M. H. & Zehnder, F. Cognition-based interventions for healthy older people and people with mild cognitive impairment. *The Cochrane Library*, https://doi.org/10.1002/14651858.CD006220.pub2 (2011).
14. Verhaeghen, P., Marcoen, A. & Goossens, L. Improving memory performance in the aged through mnemonic training: A meta-analytic study. *Psychol. Aging* **7**, 242–251 (1992).
15. Zehnder, F., Martin, M., Altgassen, M. & Clare, L. Memory training effects in old age as markers of plasticity: A meta-analysis. *Restor. Neurol. Neurosci.* **27**, 507–520 (2009).

16. Gross, A. L. *et al*. Memory training interventions for older adults: A meta-analysis. *Aging Ment. Health* **16**, 722–734 (2012).
17. Richardson, A. E., Montello, D. R. & Hegarty, M. Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Mem. Cogn.* **27**, 741–750 (1999).
18. Waller, D., Bachmann, E., Hodgson, E. & Beall, A. The HIVE: A huge immersive virtual environment for research in spatial cognition. *Behav. Res. Methods* **39**, 835–843 (2007).
19. Jimeno, A. & Puerta, A. State of the art of the virtual reality applied to design and manufacturing processes. *Int. J. Adv. Manuf. Technol.* **33**, 866–874 (2007).
20. Bronfenbrenner, U. Toward an experimental ecology of human development. *Am. Psychol.* **32**, 513–531 (1977).
21. Optale, G. *et al*. Controlling Memory Impairment in Elderly Adults Using Virtual Reality Memory Training: A Randomized Controlled Pilot Study. *Neurorehabil. Neural Repair* **24**, 348–357 (2010).
22. Borkin, M. A. *et al*. What Makes a Visualization Memorable. *IEEE Trans. Vis. Comput. Graph.* **19**, 2306–2315 (2013).
23. Christou, C. G. & Bülthoff, H. H. View Dependence in Scene Recognition after Active Learning. *Mem. Cognit.* **27**, 996–1007 (1999).
24. Meijer, F., Geudeke, B. L. & van den Broek, E. L. Navigating through Virtual Environments: Visual Realism Improves Spatial Cognition. *Cyberpsychology Behav.* **12**, 517–521 (2009).
25. Smallman, H. S. & John, M. S. Naive Realism: Misplaced Faith in Realistic Displays. *Ergon. Des.* **13**, 6–13 (2005).
26. Wilkening, J. & Fabrikant, S. I. How Do Decision Time and Realism Affect Map-Based Decision Making? In *International Conference on Spatial Information Theory* 1–19, https://doi.org/10.1007/978-3-642-23196-4_1 (2011).
27. Çöltekin, A., Francelet, R., Richter, K. F., Thoresen, J. & Fabrikant, S. I. The effects of visual realism, spatial abilities, and competition on performance in map-based route learning in men. *Cartogr. Geogr. Inf. Sci.* **45**, https://doi.org/10.1080/15230406.2017.1344569 (2017).
28. Smallman, H. S. & John, M. Naive realism: Limits of realism as a display principle. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **49**, 1564–1568 (2005).
29. Salthouse, T., Mitchell, D., Skovronek, E. & Babcock, R. Effects of adult age and working memory on reasoning and spatial abilities. *J. Exp. Psychol. Learn. Mem. Cogn.* **15**, 507–516 (1989).
30. Kirasic, K. C. Spatial cognition and behavior in young and elderly adults: implications for learning new environments. *Psychol. Aging* **6**, 10–18 (1991).
31. Burns, P. C. Navigation and the Mobility of Older Drivers. **54**, 49–55 (1999).
32. Cherry, K. E. & Park, D. C. Individual difference and contextual variables influence spatial memory in younger and older adults. *Psychol. Aging* **8**, 517–526 (1993).
33. Park, D. C. *et al*. Models of Visuospatial and Verbal Memory Across the Adult Life Span. *Psychol. Aging* **17**, 299–320 (2002).
34. Nemmi, F., Boccia, M. & Guariglia, C. Does aging affect the formation of new topographical memories? Evidence from an extensive spatial training. *Aging, Neuropsychol. Cogn.* **24**, 29–44 (2017).
35. Evans, G. W. & Pezdek, K. Cognitive mapping: knowledge of real-world distance and location information. *J. Exp. Psychol. Hum. Learn.* **6**, 13–24 (1980).
36. Moffat, S. D. & Resnick, S. M. Effects of age on virtual environment place navigation and allocentric cognitive mapping. *Behav. Neurosci.* **116**, 851–859 (2002).
37. Muffato, V., Meneghetti, C., Di Ruocco, V. & De Beni, R. When young and older adults learn a map: The influence of individual visuo-spatial factors. *Learn. Individ. Differ.* **53**, 114–121 (2017).
38. Gyselinck, V. *et al*. Considering spatial ability in virtual route learning in early aging. *Cogn. Process.* **14**, 309–316 (2013).
39. Dodson, C. S., Bawa, S. & Slotnick, S. D. Aging, source memory, and misrecollections. *J. Exp. Psychol. Learn. Mem. Cogn.* **33**, 169–181 (2007).
40. Zelinski, E. M. & Light, L. L. Young and older adults' use of context in spatial memory. *Psychol. Aging* **3**, 99–101 (1988).
41. Dodson, C. S., Koutstaal, W. & Schacter, D. L. Escape from illusion: Reducing false memories. *Trends Cogn. Sci.* **4**, 391–397 (2000).
42. Kelley, C. M. & Sahakyan, L. Memory, monitoring, and control in the attainment of memory accuracy. *J. Mem. Lang.* **48**, 704–721 (2003).
43. Dodson, C. S., Bawa, S. & Krueger, L. E. Aging, metamemory, and high-confidence errors: A misrecollection account. *Psychol. Aging* **22**, 122–133 (2007).
44. Pardilla-Delgado, E. & Payne, J. D. The Deese-Roediger-McDermott (DRM) Task: A Simple Cognitive Paradigm to Investigate False Memories in the Laboratory. *J. Vis. Exp.*, 1–10, https://doi.org/10.3791/54793 (2017).
45. Koutstaal, W. & Schacter, D. L. Gist-Based False Recognition of Pictures in Older and Younger Adults. **37**, 555–583 (1997).
46. Dodson, C. S. & Krueger, L. E. I misremember it well: why older adults are unreliable eyewitnesses. *Psychon. Bull. Rev. Rev.* **13**, 770–775 (2006).
47. Zhong, J. Y. & Moffat, S. D. Age-Related Differences in Associative Learning of Landmarks and Heading Directions in a Virtual Navigation Task. *Front. Aging Neurosci.* **8**, 3–11 (2016).
48. Estes, W. K. In *Handbook of Learning and Cognitive Processes* **5**, 271–295 (2014).
49. Hegarty, M. Dynamic visualizations and learning: Getting to the difficult questions. *Learn. Instr.* **14**, 343–351 (2004).
50. Kirschner, P. A. Cognitive Load Theory: Implications of Cognitive Load Theory on the Design of Learning. *Learing Instr.* **12**, 1–10 (2002).
51. Schnotz, W. Towards an integrated view of learning from text and visual displays. *Educ. Psychol. Rev.* **14**, 101–120 (2002).
52. Lokka, I. E. & Çöltekin, A. Remembering what we see: Designing virtual environments to improve visuo-spatial recall for navigation tasks. In *Proceedings of the 28th International Cartographic Conference* (2017).
53. Lokka, I. E. & Çöltekin, A. Toward optimizing the design of virtual environments for route learning: empirically assessing the effects of changing levels of realism on memory. *Int. J. Digit. Earth*, 1–19 (2017).
54. Waller, D. & Lippa, Y. Landmarks as beacons and associative cues: Their role in route learning. *Mem. Cognit.* **35**, 910–924 (2007).
55. Lenneberg, E. H. Color Naming, Color Recognition, Color Discrimination: A Re-Appraisal. *Percept. Mot. Skills* **12**, 375–382 (1961).
56. Mayer, R. E. & Moreno, R. Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educ. Psychol.* **38**, 43–52 (2003).
57. Çöltekin, A., Bleisch, S., Andrienko, G. & Dykes, J. Persistent challenges in geovisualization – a community perspective. *Int. J. Cartogr.* **3**, 1–25 (2017).
58. Yerkes, R. M. & Dodson, J. D. The relation of strength of stimulus to rapidity of habit-formation. *J. Comp. Neurol. Psychol.* **18**, 459–482 (1908).
59. Röser, F., Krumnack, A., Hamburger, K. & Knauff, M. A four factor model of landmark salience – A new approach. *Proc. 11th Int. Conf. Cogn. Model.* 82–87 (2012).
60. Röser, F., Hamburger, K., Krumnack, A. & Knauff, M. The Structural Salience of Landmarks: Results from an On-Line Study and a Virtual Environment Experiment. *J. Spat. Sci.* **57**, 37–50 (2012).
61. Karimpur, H., Röser, F. & Hamburger, K.Finding the Return Path: Landmark Position Effects and the Influence of Perspective. F*ront. Psychol.* 7 (2016).
62. Winter, S. Route Adaptive Selection of Salient Features. In *International Conference on Spatial Information Theory* 349–361, https://doi.org/10.1007/978-3-540-39923-0_23 (2003).
63. Richter, K.-F. & Winter, S. *Landmarks*, https://doi.org/10.1007/978-3-319-05732-3 (2014).
64. Raubal, M. & Winter, S. Enriching Wayfinding Instructions with Local Landmarks. *Proc. Second Int. Conf. Geogr. Inf. Sci.*, 243–259 (2002).

65. Klippel, A. & Winter, S. Structural Salience of Landmarks for Route Directions. In *International Conference on Spatial Information Theory* (ed. Springer Berlin Heidelberg), 347–362, https://doi.org/10.1007/11556114_22 (2005).
66. Waller, D. & Greenauer, N. The role of body-based sensory information in the acquisition of enduring spatial representations. *Psychol. Res.* **71**, 322–332 (2007).
67. Wiener, J. M., de Condappa, O., Harris, M. A. & Wolbers, T. Maladaptive bias for extrahippocampal navigation strategies in aging humans. *J. Neurosci.* **33**, 6012–7 (2013).
68. Strickrodt, M., O'Malley, M. & Wiener, J. M. This Place Looks Familiar — How Navigators Distinguish Places with Ambiguous Landmark Objects When Learning Novel Routes. *Front. Psychol.* **6** (2015).
69. Aginsky, V., Harris, C., Rensink, R. & Beusmans, J. Two strategies for learning a route in a driving simulator. *J. Environ. Psychol.* **17**, 317–331 (1997).
70. Claramunt, C. & Winter, S. Structural Salience of Elements of the City. *Environ. Plan. B Plan. Des.* **34**, 1030–1050 (2007).
71. Merriman, N. A. *et al.* Crowded environments reduce spatial memory in older but not younger adults. *Psychol. Res.* **82**, 1–22, https://doi.org/10.1007/s00426-016-0819-5 (2016).
72. Rosenbaum, R. S., Winocur, G., Binns, M. A. & Moscovitch, M. Remote spatial memory in aging: all is not lost. *Front Aging Neurosci.* **4**, 1–10 (2012).
73. Moffat, S. D., Elkins, W. & Resnick, S. M. Age differences in the neural systems supporting human allocentric spatial navigation. *Neurobiol. Aging* **27**, 965–972 (2006).
74. Wiener, J. M., Kmecova, H. & de Condappa, O. Route repetition and route retracing: Effects of cognitive aging. *Front. Aging Neurosci.* **4**, 1–7 (2012).
75. Janzen, G. Memory for object location and route direction in virtual large-scale space. *Q. J. Exp. Psychol.* **59**, 493–508 (2006).
76. Charness, N. Visual short-term memory and aging in chess players. *J. Gerontol.* **36**, 615–9 (1981).
77. Fandakova, Y. *et al.* Age differences in false memory: The importance of retrieval monitoring processes and their modulation by memory quality. *Psychol. Aging* **33**, 119–133 (2018).
78. MacDonald, S. W. S., Stigsdotter-Neely, A., Derwinger, A. & Backman, L. Rate of acquisition, adult age, and basic cognitive abilities predict forgetting: New views on a classic problem. *J. Exp. Psychol. Gen.* **135**, 368–390 (2006).
79. Old, S. R. & Naveh-Benjamin, M. Differential Effects of Age on Item and Associative Measures of Memory: A Meta-Analysis. *Psychol. Aging* **23**, 104–118 (2008).
80. Sporer, S. L., Penrod, S., Read, D. & Cutler, B. Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychol. Bull.* **118**, 315–327 (1995).
81. Smallman, H. S. & Cook, M. B. Naïve Realism: Folk fallacies in the design and use of visual displays. *Top. Cogn. Sci.* **3**, 579–608 (2011).
82. O'Bryant, S. E. *et al.* Detecting Dementia with the Mini-Mental State Examination (MMSE) in Highly Educated Individuals. *Archives of neurology* **65**, 963–967 (2008).
83. Sullivan, C. O. & Newell, F. N. Crowded environments reduce spatial memory in older but not younger adults. *Psychological research* **82**, 407–428, https://doi.org/10.1007/s00426-016-0819-5 (2016).
84. Sekuler, R., Kahana, M. J., McLaughlin, C., Golomb, J. & Wingfield, A. Preservation of Episodic Visual Recognition Memory in Aging. *Exp. Aging Res.* **31**, 1–13 (2005).

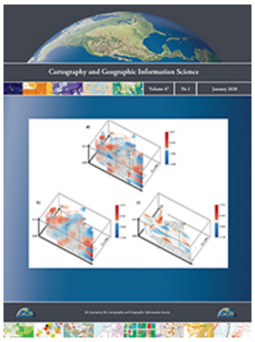
Acknowledgements
This work was supported by the Swiss National Science Foundation (Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung) [grant number 200021_149670] (SNSF project VISDOM). The authors want to thank the URPP (University Research Priority Program) "Dynamics of Healthy Aging" from the University of Zurich for recruiting participants for the experiment and Robert Weibel for his valuable input and feedback for the creation of this paper.

Author Contributions
Authors contributed according to their competences and interests. I.L. and A.C. built the concept. I.L. implemented and performed the experiment; A.C., S.I.F. and C.R. consulted the experiment. I.L. analyzed the data; A.C. and J.W. consulted the analysis. I.L. and A.C. drafted the manuscript; S.I.F., J.W., C.R. edited and revised the manuscript; All authors approved the final version of the manuscript.

Additional Information
Competing Interests: The authors declare no competing interests.

Publication III

Lokka, I.E., & Çöltekin, A., (2020). Perspective switch and spatial knowledge acquisition: effects of age, mental rotation ability and visuospatial memory capacity on route learning in virtual environments with different levels of realism. *Cartography and Geographic Information Science*, 47(1), 14-27. https://doi.org/10.1080/15230406.2019.1595151



Taylor & Francis
Taylor & Francis Group

Cartography and Geographic Information Science

ISSN: 1523-0406 (Print) 1545-0465 (Online) Journal homepage: https://www.tandfonline.com/loi/tcag20

Perspective switch and spatial knowledge acquisition: effects of age, mental rotation ability and visuospatial memory capacity on route learning in virtual environments with different levels of realism

Ismini E. Lokka & Arzu Çöltekin

To cite this article: Ismini E. Lokka & Arzu Çöltekin (2020) Perspective switch and spatial knowledge acquisition: effects of age, mental rotation ability and visuospatial memory capacity on route learning in virtual environments with different levels of realism, Cartography and Geographic Information Science, 47:1, 14-27, DOI: 10.1080/15230406.2019.1595151

To link to this article: https://doi.org/10.1080/15230406.2019.1595151

Published online: 23 Apr 2019.

CARTOGRAPHY AND GEOGRAPHIC INFORMATION SCIENCE
2020, VOL. 47, NO. 1, 14–27
https://doi.org/10.1080/15230406.2019.1595151

ARTICLE

Perspective switch and spatial knowledge acquisition: effects of age, mental rotation ability and visuospatial memory capacity on route learning in virtual environments with different levels of realism

Ismini E. Lokka [a] and Arzu Çöltekin [a,b]

[a]Department of Geography, University of Zurich, Zurich, Switzerland; [b]Institute for Interactive Technologies, University of Applied Sciences and Arts Northwestern Switzerland, Brugg-Windish, Switzerland

ABSTRACT

We report on a study in which we examine if the visual design of virtual environments (VEs) affects visuospatial knowledge acquisition in younger and older adults with varying cognitive abilities in the context of navigational learning, *specifically* when a perspective switch is involved. Perspective switch between first-person and aerial-views is an important and commonly executed task in navigation; and it is a special case in studying the effects of aging on navigational performance as well, because, reportedly, it is particularly harder for older people. In a controlled experiment, our participants learned a route in first-person view VE, and reproduced what they learned in an aerial-perspective view in immediate and delayed recall stages. To examine the effects of (and interactions between) multiple factors involved in the experiment in relation to the given task, we provide an in-depth investigation of group differences in spatial knowledge acquisition when a perspective switch is required based on age, mental rotation abilities, and visuospatial memory capacity with three VE designs that differ in levels of realism. Our findings based on the recall accuracy of 81 (42 younger, 39 older) participants in sketching tasks demonstrate significant differences across VE types, overall, in favor of our custom-designed VE in this demanding task. Furthermore, we demonstrate that age and visuospatial memory abilities are strong moderating factors, explicitly in this sketching task that requires a perspective switch, irrespective of VE types.

ARTICLE HISTORY
Received 15 October 2018
Accepted 11 March 2019

KEYWORDS
Age; spatial ability; memory capacity; spatial knowledge acquisition; mental rotation; perspective switch; delayed recall

Introduction

When people follow a route, it is commonly assumed that the spatial structure of the environment is encoded in the human mind, and translated into spatial knowledge (Golledge, Dougherty, & Scott, 1995). This spatial knowledge acquisition can be affected by the visuospatial information that a person experiences (e.g. *what* is in the scene, and *what is where* in the scene), as well as the individual differences such as prior experience levels, memory capacity, and spatial abilities. While the effect of some of these factors are well-documented in other contexts, their combined effects and interactions with age and long-term information retention in the context of route learning with custom-designed VEs are not well understood. More specifically, we do not know how aging, tasks that require perspective switching in spatial knowledge acquisition, and abilities (such as mental rotation ability and visuospatial memory capacity) interact with differently designed virtual environments designed for route learning. Perspective switching in navigational tasks is known to be especially difficult for older adults (Herman & Coyne, 1980; Inagaki et al., 2002), but would this be different if a route learning environment is deliberately modified in terms of visual content?

To address the gap stated above, in this paper we examine: 1) whether we can improve spatial knowledge acquisition linked to *perspective switching tasks* in the context of route learning by optimizing a VE's visual design, 2) how age, mental rotation abilities, and visuospatial memory capacity interact with different VE designs in spatial knowledge acquisition in perspective switching tasks during (virtual) route learning. We hypothesize that with a deliberate visualization design (i.e. manipulating the visual scene content with the intention to assist spatial knowledge acquisition in general), we can improve the rate of acquired spatial knowledge. In previous analyses, we have shown that such deliberate designs *do* assist people of different age groups in *tasks that require visuospatial information recall* (Lokka & Çöltekin, 2017b; Lokka, Çöltekin, Wiener, Fabrikant, & Röcke, 2018). Differently to our previous work, in this paper we examine if the observed effects persist in *tasks that require perspective switching* across different age ranges and *also*

CONTACT Ismini E. Lokka ismini.lokka@uzh.ch

ability groups, and we believe age and ability will act as moderating factors irrespective of visualization design. Below we provide a concise literature review that our hypotheses are built upon.

Understanding and measuring spatial knowledge acquisition

In general, acquiring spatial knowledge from a first-person view navigational experience is a non-trivial task, as it requires constantly filtering the relevant information from a plethora of visuospatial input. Spatial knowledge acquisition becomes even more complex if the experienced first-person view must be translated to a top-down aerial view. If a person is asked to produce a 2D sketch of the route they just walked, they must perform a mental rotation from the first-person perspective (a "street view") to an aerial perspective (a top-down view), creating a structural layout of the navigated space in their mind (Thorndyke & Hayes-Roth, 1982). The complexity of mentally transforming the first-person 3D experiences (i.e. the route knowledge) to metric and survey knowledge has been demonstrated in previous studies (Golledge et al., 1995). The reverse, i.e. transformations from a 2D map view to the "real" 3D world has been shown to be non-trivial too (Kiefer, Giannopoulos, & Raubal, 2014). Importantly, it has also been shown that the complexity of the mental transformation during perspective switch in the context of route learning differs widely among individuals (Ishikawa & Montello, 2006). Besides the perspective shift in the visual experience and individual differences, the nature of the task is also important. For tasks requiring verbal descriptions of a learned layout, people who acquired spatial knowledge through direct exposure (from a first-person view) could provide descriptions *only* based on route knowledge, while people who learned from a map (from a top-down view) could provide descriptions *both* based on route and survey knowledge (Taylor & Tversky, 1996). These findings support the position that the experiences based on direct exposure to the environment provides limited assistance in acquiring survey knowledge. Shelton and McNamara (2004) provided further evidence that perspective switching (or shifting) introduces complications. In their experiment, keeping the "test view" (i.e. when the participant needs to recognize the layout) similar to the view during encoding was beneficial in recognition speed (Shelton & McNamara, 2004). Another group of researchers compared two study groups in tasks regarding survey knowledge acquisition, and also concluded that acquiring survey knowledge requires more cognitive effort than acquiring route knowledge (Van Asselen, Fritschy, & Postma, 2006). In summary, performing additional mental rotations (either because of changes in orientation or perspective) appears to have cognitive costs.

Spatial knowledge acquisition is often evaluated by measuring the success rates in various tasks, such as pointing tasks (i.e. judgement of relative direction), distance estimations, identification of the shortest route to a target, and producing sketch maps of the learned route among others (Wang, 2017). Use of sketch maps have received criticism due to (i) the variability of the tested environments, (ii) the subjectivity of the evaluation of accuracy, and (iii) the inability to control the participant's experience (Ishikawa & Montello, 2006). Nevertheless, sketch maps remain as a common metric in spatial knowledge acquisition studies (Appleyard, 1970; Billinghurst & Weghorst, 1995; Blades, 1990; Curtis, 2016; Lynch, 1960; Witmer, Sadowski, & Finkelstein, 2002), possibly because they are a common, and fairly intuitive, way to describe a route. Proposing an objective evaluation scheme for sketch maps is a non-trivial task, and there is a lack of a clearly defined methodology (Billinghurst & Weghorst, 1995). However, various "rules of thumb" can be obtained from the related work. Some important factors to consider in evaluating a sketch map are (Ladd, 1970; Moore, 1976): number of landmarks, number of streets (Anacta, Wang, & Schwering, 2014), object classes (Billinghurst & Weghorst, 1995), and relative object positioning (Billinghurst & Weghorst, 1995). Ideally these factors are presented on a sketch map to "sufficiently" represent the experienced route, and from there one can infer that the individual was successful in selecting relevant information, and in mentally processing the perspective switch.

The examples reviewed so far stress the difficulty of the acquisition of metric or survey knowledge from a top-down perspective, especially when the encoding occurs from a first-person perspective view. Aside from the factors related to the viewing perspective, a number of other factors affect spatial knowledge acquisition too. For example, it has been shown that if the learning is *intentional* or *incidental*, it affects spatial knowledge acquisition (Van Asselen et al., 2006). Importantly, it is well-understood that participant characteristics (i.e. individual and group differences, such as expertise, spatial abilities, and age) play a significant role in spatial knowledge acquisition.

Effects of aging on spatial knowledge acquisition

In general, a wide range of abilities, skills, and attitudes play a role in successful spatial learning (Weisberg & Newcombe, 2016; Wolbers & Hegarty, 2010); and some of these skills can be assessed using standardized tests

(Ekstrom, French, Harman, & Dermen, 1976; Vandenberg & Kuse, 1978). In spatial knowledge acquisition, differences in individuals' spatial abilities and memory capacity are considered especially relevant (Çöltekin, Francelet, Richter, Thoresen, & Fabrikant, 2018; Hegarty, Montello, Richardson, Ishikawa, & Lovelace, 2006; Hurlebaus, Basten, Mallot, & Wiener, 2008; Ishikawa & Montello, 2006; Montello, 1998; Muffato, Meneghetti, & De Beni, 2016; Muffato, Meneghetti, Di Ruocco, & De Beni, 2017; Wolbers & Hegarty, 2010). Importantly, these abilities are affected by age-related cognitive decline. Specifically, it is well-known that aging negatively correlates with navigational learning (Moffat, 2001; Wiener, Kmecova, & de Condappa, 2012; Wolbers, Dudchenko, & Wood, 2014) and performance in spatial memory (Richmond, Sargent, Flores, & Zacks, 2018). This appears to be true especially in allocentric tasks (Fricke & Bock, 2018). Thus, aging is necessary to consider as a factor in spatial knowledge acquisition studies.

It appears that switching from an egocentric (e.g. first-person view) perspective to an allocentric (e.g. top-down view) one is a "weak spot" particularly for older people. Several studies featuring tasks that require a perspective switch demonstrated that older people commit more mental rotation errors than younger people (e.g. Herman & Coyne, 1980; Inagaki et al., 2002). Furthermore, when there are many orientation changes throughout the learned route, older adults are less effective in environmental learning, irrespective of the perspective in which the environment is experienced (Yamamoto, Fox, Boys, & Ord, 2018). In Yamamoto et al.'s (2018) study, learning a route from an egocentric representation impaired survey knowledge acquisition in older people, whereas learning from an allocentric representation did not, that is, results were comparable to that of younger people (Yamamoto & DeGirolamo, 2012). A recent study examining sketch map accuracy added further nuance to this finding: After learning from an allocentric representation, older adults' sketches were less accurate than those of younger adults, if *missed locations of landmarks* were used as a measure of accuracy (Muffato et al., 2017). However, authors demonstrated that when only *number of landmarks* is used as a measure of accuracy, age differences disappeared. In Muffato et al.'s (2017) study, visuospatial working memory capacity correlated with success irrespective of age (Muffato et al., 2017).

The examples reviewed above are only a small portion of the vast literature that demonstrates how individual and group differences can affect spatial knowledge acquisition, especially when there are orientation or perspective shifts during learning. The fact that spatial (and especially survey) knowledge acquisition is more difficult for some people than others highlights the necessity to address individual and group differences while designing visualizations (including VEs), so that visual displays, such as VEs, facilitate better spatial learning for all. The concept of *designing for all* in the geovisualization literature is well investigated, especially for maps (Reichenbacher, 2001). Maps that are "designed for all" respond to any accessibility issues, and accordingly adapt to user needs. Such maps (and visualizations) can be personalized, and/or optimized for group differences based on age and expertise (Nivala & Sarjakoski, 2005). By ensuring that designs are *accessible* to all, is reasonable to assume that they are *improved* for everyone, irrespective of abilities.

Virtual environments for route learning in navigational tasks

VEs have long been used for studying navigation and spatial cognition. As opposed to the real world, VEs provide safe and controlled environments, thus, we can examine navigational behavior (and associated spatial learning) in response to a specific variable of interest, without other variables confounding. However, it is important to remember that the way these VEs are visually designed—for example, the amount, the quality, and the location of provided information within a VE (Lokka & Çöltekin, 2017a)—can have a strong impact on the spatial learning performance. Through a well-informed visualization design, one might be able to improve knowledge acquisition from VEs for everyone. Such an improvement might be especially relevant for those with lower abilities (e.g. people with lower visuospatial abilities, memory capacity, or older adults). It has been previously demonstrated that differently designed VEs facilitate spatial learning differently. More specifically, custom-designed VEs can improve short- and long-term recall performance in visuospatial tasks (Lokka & Çöltekin, 2017b). This effect is also relevant for older adults. With such custom designed VEs, older people improve their accuracy in visuospatial knowledge acquisition, and calibrate their confidence in tasks that *retain the viewing perspective* (Lokka et al., 2018). A well-considered adjustment of the visualization design is known to reduce cognitive load (Sweller, 1988), and we believe such an adjustment will also have direct effects on tasks that require *perspective switching*. A well-considered adjustment, for example, could include levels of realism in the context of working with VEs. High levels of visual realism have been shown to negatively correlate with cognitive load, and it can impair the rates of spatial knowledge acquisition, especially for people with lower spatial abilities (Lokka & Çöltekin, 2017b; Lokka et al., 2018). Despite the cognitive load it seems to introduce, people overwhelmingly *prefer* visually realistic displays to more abstract ones, possibly because of their resemblance to the real world and the associated feeling of familiarity (Çöltekin et al., 2017; Lokka et al., 2018; Smallman & John, 2005).

Because photo-realistic representations are popular, yet they might impair performance in tasks that are demanding on working memory; we believe a visualization solution that balances between preference and performance is using realistic photo-textures *selectively* in spatial learning tasks. Doing so highlights the important information that aids encoding relevant visuospatial features in the scene, yet keeps "enough" realism to simulate the *sense of place*, and to provide a reference (or an anchor) to the real world. When selectively showing the photo-textured features, an important design consideration is *where* these photo-textured features should be located. Besides their location, saliency of landmarks can be defined by semantic, visual, and structural elements, and each of these are important for the *attractiveness of landmarks* (Raubal & Winter, 2002). In this paper we focus on the *locations* of the highlighted features, because as soon as we highlight selected features, they serve as landmarks, and landmarks are important facilitators of spatial learning (Richter & Winter, 2014; Röser, Hamburger, Krumnack, & Knauff, 2012; Winter, Raubal, & Nothegger, 2005). Thus, in addition to adjusting levels of realism against cognitive load; we believe the photo-textured features (i.e. "landmarks") should be positioned in locations where people would (and would need to) pay attention, based on previous findings (Röser et al., 2012).

Hypotheses

We implemented a visualization design solution that we believe would balance levels of realism for optimum route learning performance while remaining attractive to the users and called this solution "MixedVE". Previously, we tested the MixedVE against the two baseline VEs with younger participants in a variety of tasks that involve visual, spatial, and visuospatial recall of information and confirmed its value to our younger age group (Lokka & Çöltekin, 2017b). Furthermore, we already investigated how the MixedVE facilitates the recall of visuospatial information for older adults for visuospatial tasks, i.e. we examined the recall rates of both age groups in a task that did *not* involve perspective switching (Lokka et al., 2018). To acquire a holistic understanding of the effect of the MixedVE for navigational tasks, we now further investigate the MixedVE's performance in a *perspective switching task*; we compare participants' spatial knowledge acquisition in a route learning task where participants need to sketch a 2D top-down representation of a route they learned in a first-person perspective VE. We conducted the tests with the MixedVE against two other VEs: An abstract one with no visual cues (AbstractVE), and a fully photo-textured one that resembles a realistic environment (RealisticVE). We examine if the benefits offered by the MixedVE, specifically *for this task type*, transcends individual differences based on age, mental rotation ability, and visuospatial memory capacity. We believe that the MixedVE will provide benefits over the two alternatives we tested.

Comparing the MixedVE to the Abstract and Realistic VEs as described above (and illustrated in Figure 1), we specifically investigate: 1) whether the MixedVE assists in spatial knowledge acquisition in tasks that involve perspective switching more than the Abstract and Realistic VEs (measured in active sketching tasks); 2) whether the observed differences in the accuracy of acquired spatial knowledge in tasks that involve perspective switching (if any) are explained by differences in age, mental rotation (MRT) abilities, or visuospatial memory (VSM) capacity and how these interact with visualization types (the three VEs); and, 3) whether MRT and VSM tests predict the successful acquisition of the spatial knowledge in tasks that involve switching perspectives, especially in relation to different visualization types. Based on the literature, for each question framed above, we hypothesize the following:

(1) The MixedVE will facilitate better spatial knowledge acquisition in tasks that involve perspective switching (i.e. accuracy in sketching) than the other two VEs.
(2) Younger participants will produce more accurate sketches than older participants, irrespective of VE type; and, participants with higher MRT/VSM scores will outperform the participants with lower MRT/VSM scores irrespective of age or VE type.
(3) Irrespective of age, participants with higher VSM will outperform the participants with lower VSM in producing accurate sketches, particularly with the MixedVE and RealisticVE, as these provide potentially helpful photographic visual cues; whereas MRT will be most relevant to Abstract VE because this visualization type contains no (photographic) visual cues.

Methods

To test our hypotheses, we conducted a controlled laboratory experiment. In our mixed-factorial design,

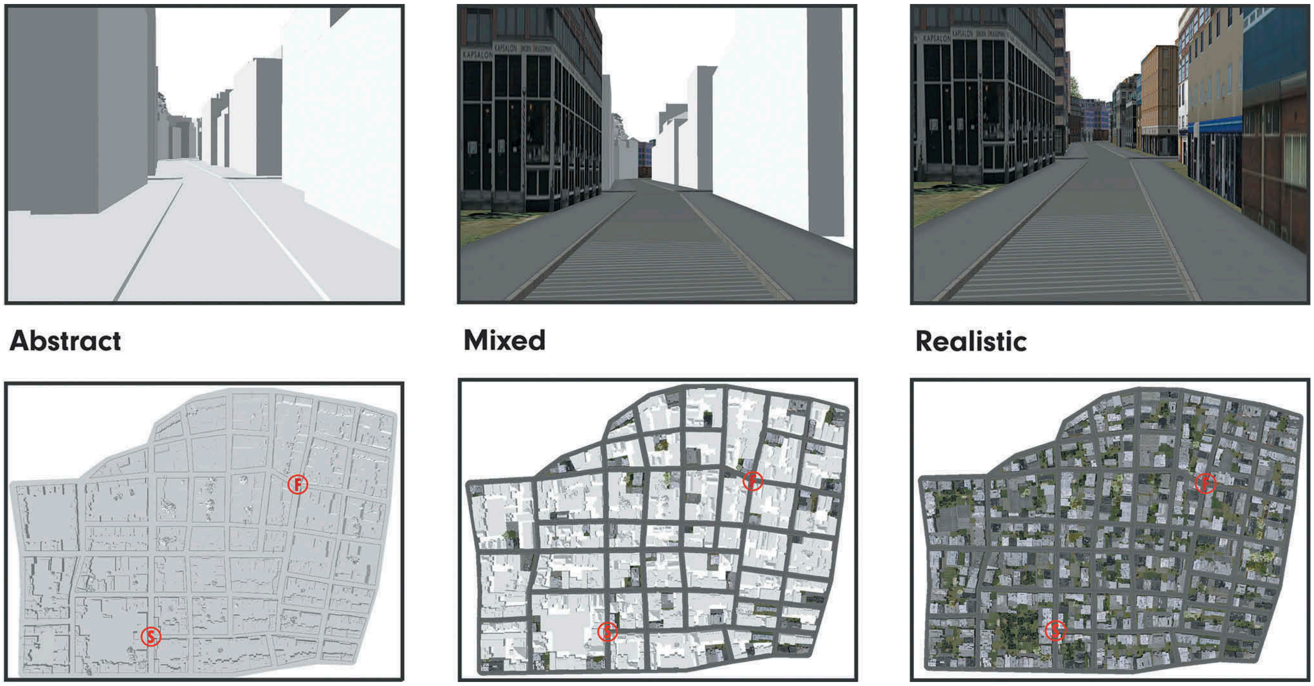


Figure 1. Top: The three experimental conditions in which all participants experienced a video drive-through: the abstract, mixed, and realistic VEs (left to right). Down: The 2D map views of each VE, on which participants sketched the respective routes they experienced in each VE from the memory. Start and end points were marked on the screenshots.

independent variables were: *age* as a between-subject factor (younger and older) and *visualization type* (Abstract, Mixed, Realistic VEs) as a within-subject factor. Participants were explicitly asked to learn a given route presented to them as a video. We have previously published two papers that are based on different data that are collected from the same participants, based on different experimental tasks, answering different research questions (Lokka & Çöltekin, 2017b; Lokka et al., 2018). In the experiment, participants performed a set of visuospatial recall tasks analyzed in Lokka and Çöltekin (2017b), along with an active *sketching task*. In this paper, we *only* focus on the latter task, which has not been analyzed before. During this active sketching task, the participants drew the route they experienced in the VEs on printed top-down screen shots of each VE (thus the visual cues in the baseline map differed according to the VE type, see Figure 1). We selected such a task for experimental control: The visual stimuli the participants experienced during the *encoding* were kept identical (i.e. everyone experienced the same visual scenes, thus did not have different landmarks), and the same was true for the visual stimuli at the *decoding* stage (i.e. 2D map views provided for the active sketching were identical). By keeping the visual variability in check, we ensure the comparability of the success rates we measure per task. Participants drew the sketches twice from the memory; thus, we measured *immediate recall* success right after they experienced the VEs, and *delayed recall* success a week later. The sketches were evaluated for their accuracy and completeness based on a scoring scheme, and this was our main dependent variable.

Participants

A total of 81 volunteers took part in the experiment; 42 of them were younger (23 female, 20–30 yrs.) and 39 older (17 female, 65–75 yrs.). Younger participants were recruited by word of mouth, while the older participants were recruited via the participant pool of the University Research Priority Program "Dynamics of Healthy Aging" of the University of Zurich. We screened all participants for mild cognitive impairments as an inclusion/exclusion criteria; that is, they were included in the experiment if they achieved a score of 27 and above on the Mini Mental State Examination (MMSE) (O'Bryant et al., 2008).

Materials

Stimuli

We conducted the experiment in a 3D visualization lab (Department of Geography, University of Zurich). VEs were shown on a large rear-projection display (2438mm x 1829mm), at 2.2m distance from the participants. The VEs featured a fictitious 3D city with buildings similar in style, shape and color to control for their possible effects on the memory because of their distinctiveness. We designed the angles at the intersection and turn points to control for visibility of features. We then selected two routes of equal length, with comparable visual information, and equal amount of "turns" (3 left, 3 right, 1 intersection continuing straight). VEs were shown as passive videos, deliberately avoiding any interaction from the participants, to make sure that participants were all exposed to the same information, for equal durations at equal speed. Routes were shown in a "driving simulation" at a constant eye level, with a fixed speed (30 km/h). We presented our fictitious city in three different visualization designs (Figure 1): the AbstractVE with no colors and no photo-textures, the RealisticVE with full visual realism with color photo-textures, and the MixedVE with a "combination" of the two in terms of realism. In the MixedVE, only the buildings at the intersection points (at critical locations) and toward the direction of turn were highlighted using color photo-textures. We also counterbalanced the content of the photo-textures for distinctness and memorability (Lokka & Çöltekin, 2017a). Furthermore, we highlighted the structural network (i.e. the road with the pavement) in the MixedVE, as this might be helpful in forming spatial knowledge.

Assessment of individual differences

We used two standardized tests to assess participants' spatial abilities and visuospatial memory capacity: 1) *Mental Rotation Test (MRT)*. This test requires correctly discriminating rotated 3D objects from foils based on a reference shape (Vandenberg & Kuse, 1978). The MRT has been used as a measure for identifying differences in spatial abilities (Meneghetti, Borella, Pastore, & De Beni, 2014; Meneghetti, Muffato, Varotto, & De Beni, 2017; Muffato, Della Giustina, Meneghetti, & De Beni, 2015; Muffato et al., 2017), even though there are arguments for the dissociation between object-based and egocentric spatial transformations (Hegarty & Waller, 2004, p. 2) *Visuospatial Memory Test (VSM)*. This test measures visuospatial memory capacity based on a 2D city plan. Participants study a city plan that contains 12 visually different buildings, and later need to place these buildings in their correct location on a layout that does not contain buildings (Ekstrom et al., 1976).

Experimental task

Participants were told that someone was driving them to their destination, but they would have to re-take this route again on their own later, thus should memorize the route to the best of their ability. Thus, the learning was intentional in the learning phase. At the response phase, similar to Krüger, Aslan, and Zimmer (2004) study, participants were provided with a printed 2D map of the area (top-down screenshots from each VE, on an A4 sheet, which were clear and legible), on which the position of the start and end points were marked (Figure 1). Participants were given the map with the north orientation (the initial orientation direction), and marked the route from memory (which they experienced in each VE during the experiment) on the given maps using a pen. This task, thus, measured spatial knowledge that requires a mental transformation from the first-person perspective to an aerial (top-down) perspective for the 2D sketch. Participants' initial orientation (heading direction) in the sketching task was the same as in the VEs (Shelton & McNamara, 2001).

Procedure

Upon arrival at the visualization lab, participants read and signed a written consent form. We then introduced the setup and the experimental process, and immediately after, we began with the main experiment.

We displayed the scenario on the screen and instructed the participants to memorize the route(s) just before they experienced the virtual routes. Then they (passively) watched the videos of the two routes in all three environments (total six videos). Each video was shown only once. We controlled for the order of presentation of the environments with a Latin Square design. After experiencing all six videos, participants were asked to solve visuospatial tasks (Lokka & Çöltekin, 2017b; Lokka et al., 2018) and then sketch the followed route of the six walkthroughs in the order they experienced them. After a week, they returned to the lab and performed the standardized tests, along with the same sketching task of all six walkthroughs from the memory, without watching the videos again.

Results

Participants' spatial knowledge acquisition was evaluated based on the accuracy and completeness of the sketches they drew (Figure 2 illustrates some examples). Specifically, sketches were evaluated for accuracy and completeness based on the following criteria: 1) number of turns (total, right and left), 2) correct direction of heading at the starting point, 3) correct direction of arriving at the end point, 4) correct direction of turn at each intersection point, 5) sequential order of turns (route patterns).

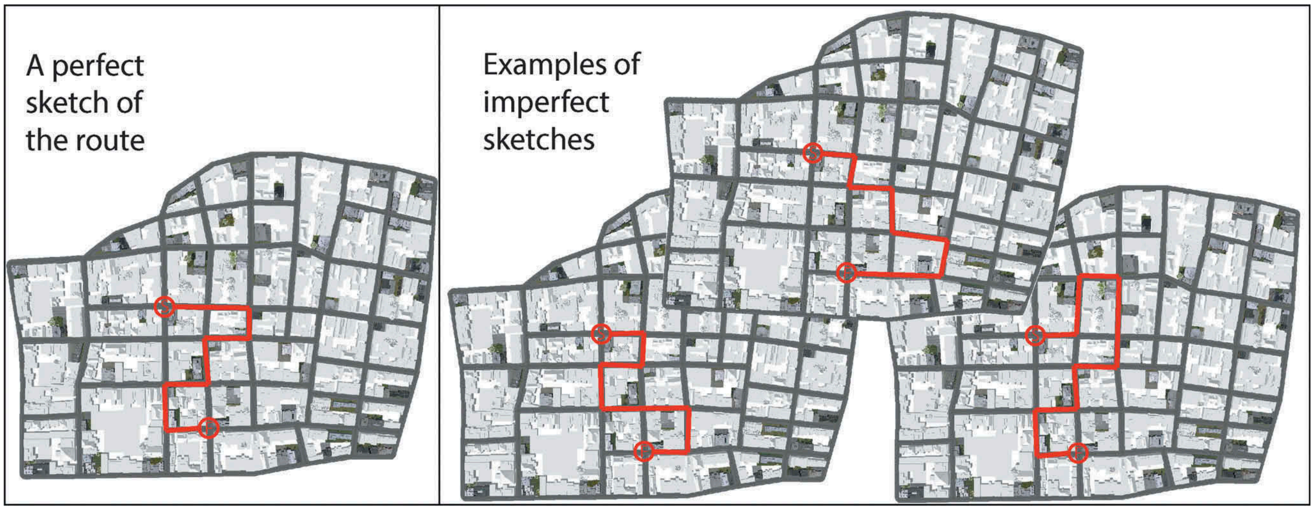


Figure 2. Example results from the sketching task. Left: A fully correct solution for one of the routes achieving a 100% score. Right: Three examples of participants' sketches of the same route with errors in terms of: number of turns, direction of turns, and sequential order of turns.

Table 1. Correlation matrix for the examined factors based on participant performance in the sketching task in both stages in all visualization conditions.

	MRT	VSM	Age
Individual and group differences			
VSM	.38***	-	
Age	−.45***	−.62***	-
Visualization conditions at the immediate (i) recall stage			
Abstract VE (i)	.26*	.33**	−.40***
Mixed VE (i)	.27*	.30**	−.45***
Realistic VE (i)	.20	.24*	−.29**
Visualization conditions at the delayed (d) recall stage			
Abstract VE (d)	.25*	.32**	−.39***
Mixed VE (d)	.22	.39***	−.44***
Realistic VE (d)	.24*	.34**	−.39***

*** $p < .001$, ** $p < .01$, * $p < .05$.

Below we begin by presenting correlations between our main variables (MRT, VSM, age, visualizations, recall stage) to provide the overall findings; and then extend the analysis to each ability test (MRT and VSM) to identify main effects and interactions in depth.

After obtaining the scores for each participant's sketches in immediate and delayed recall stages; to get an overview of how all factors in the experiment interacted, we analyzed correlations between MRT, VSM, and age groups; both with each other and with the sketching success based on all VEs (Table 1).

The effect of spatial abilities on the sketching success

To get a deeper understanding of the relationship between MRT and VSM scores and task success, we conducted two separate analyses of variance. Similar to Meneghetti, Gyselinck, Pazzaglia, and De Beni (2009) and Pazzaglia and De Beni (2006) who split their sample into high and low abilities, we grouped the participants based on a median split (Iacobucci, Posavac, Kardes, Schneider, & Popovich, 2015) (excluding the median values) into two groups for each test. The *High MRT* (n = 36) vs. *Low MRT* (n = 36), *High VSM* (n = 39) vs. *Low VSM* (n = 38) groups were treated separately in a mixed-design ANOVA, where we kept *age, recall stage,* and *visualization type* also as independent variables in both analyses.

MRT

Figure 3 shows the overall differences based on age, recall stage, MRT abilities, visualization type, and the significant interactions based on the differences in participants' MRT scores.

We see that, irrespective of their MRT scores, younger participants outperform the older (Figure 3(a)); participants are overall more successful at the sketching task in the immediate recall stage than in the delayed (Figure 3(b)); and they are more successful in sketching task based on what they recalled from the MixedVE than the other two VEs (Figure 3(d)). We also see a clear pattern that the High MRT group outperforms the Low MRT group, irrespective of other factors (Figure 3(c)). A 2 (age) x 2 (recall stage) x 2 (MRT score) x 3 (visualization type) mixed-design ANOVA revealed that all observed differences in the sketching performance for all four independent

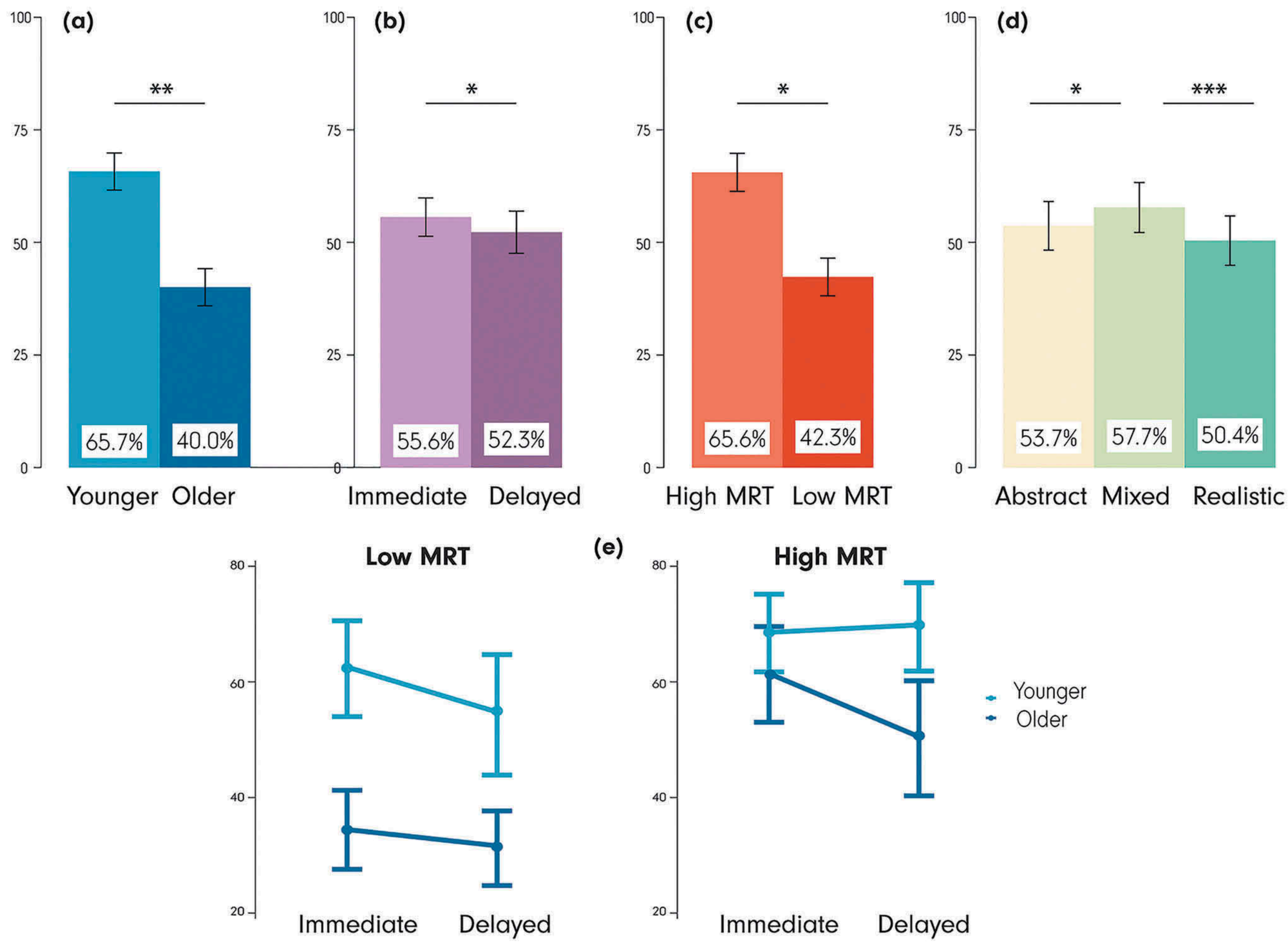


Figure 3. Main effects of (a) *age*, (b) *recall stage*, (c) *MRT score*, and (d) *visualization type* on sketching task and (e) significant interactions of *MRT ability x age x recall stage*.
*** $p < .001$, ** $p < .01$, * $p < .05$. Error bars: SEM.

variables are statistically significant; Figure 3(a)) *age* F(1, 68) = 7.81, $p < .01$, $\eta_p^2 = .08$, Figure 3(b)) *recall stage* F(1, 68) = 5.97, $p < .05$, $\eta_p^2 = .01$, Figure 3(c)) *MRT score* F(1, 68) = 5.79, $p < .05$, $\eta_p^2 = .06$, Figure 3(d)) *visualization type* F(2, 136) = 12.16, $p < .001$, $\eta_p^2 = .01$. Because there are three visualization conditions, we conducted pairwise comparisons, which revealed significant differences in participants' sketching performance (thus, "recall accuracy", as the sketches were drawn from the memory) between the three VEs: participants' overall average sketching performance was higher with the MixedVE than with the Abstract ($p < .05$, d = 0.12) and the Realistic VEs ($p < .001$, d = 0.22). Importantly, among the interactions between the four independent variables, only *age x MRT x recall stage* F(1, 68) = 4.31, $p < .05$, $\eta_p^2 = .01$ was significant (Figure 3(e)).

Because of this interaction effect, we examined the *forgetting rates* (differences in sketching performance between immediate and delayed recall stages) for older and younger participants. The analyses revealed no significant differences for the low MRT participants between the two age groups (i.e. they forget, or retain, a similar amount of information), whereas for the high MRT groups, the forgetting rate of the older participants is significantly higher than that of the younger participants (t(46.71) = −2.61, $p < .05$, r = .36).

VSM

Figure 4 shows the results of the analyses based on the VSM sample. Note that the sample size for the VSM analyses is slightly different from sample size in the MRT analyses, because we removed the median scores from the pool, and the number of people who achieved the median score was different for the two tests (participants achieving median in: MRT: $n = 9$, VSM: $n = 4$). Despite this slight difference, we see that the results overall display a similar pattern (Figure 4).

First, descriptive statistics suggest that the younger participants outperform the older (Figure 4(a)); sketching task seems to be overall easier in the immediate recall stage than

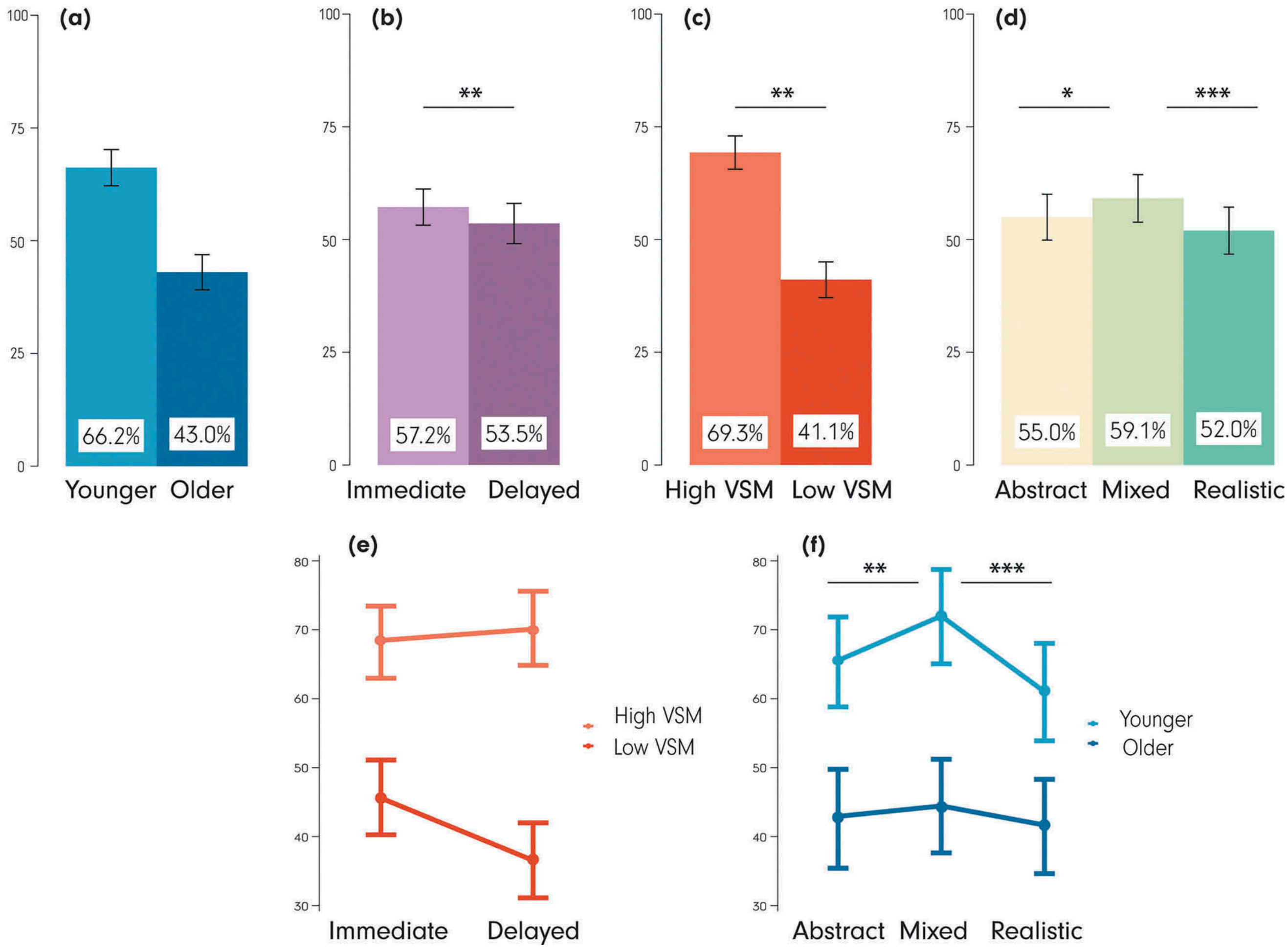


Figure 4. Main effects of (a) *age*, (b) *recall stage*, (c) *VSM scores*, and (d) *visualization type* on sketching task and significant interactions of (e) *VSM ability x recall stage* and (f) *age x visualization.*

*** $p < .001$, ** $p < .01$, * $p < .05$. Error bars: SEM.

in the delayed recall stage irrespective of other factors (Figure 4(b)), and again, MixedVE facilitates a higher success in sketching than the other two VEs (Figure 4(d)). High VSM group too, outperforms the low VSM irrespective of other factors (Figure 4(c)).

A 2 (age) x 2 (recall stage) x 2 (VSM score) x 3 (visualization type) mixed-design ANOVA revealed significant differences in the sketching performance for three out of four independent variables. Differences in sketching success based on *age* was not significant (F(1, 73) = 3.10, $p > .05$, $\eta_p^2 = .03$) (Figure 4(a)); whereas *recall stage* F(1, 73) = 8.17, $p < .01$, $\eta_p^2 = .01$ (Figure 4(b)); *VSM score* F(1, 73) = 10.79, $p < .01$, $\eta_p^2 = .11$ (Figure 4(c)); and *visualization type* F(2, 146) = 9.05, $p < .001$, $\eta_p^2 = .01$ (Figure 4(d)) led to statistically significant differences. For the visualization type, we again conducted pairwise comparisons, and observed significant differences in participants' sketching scores based on the three VEs. Specifically, again, sketching score was higher with the MixedVE than with the Abstract ($p < .05$, d = 0.13) and the Realistic VEs ($p < .001$, d = 0.22). Among the interactions between the four independent variables, *VSM score x recall stage* F(1, 73) = 7.00, $p < .01$, $\eta_p^2 = .01$ and *age x visualization* F(2, 146) = 3.32, $p < .05$, $\eta_p^2 = .00$ were significant (Figure 4(e)). In *age x visualization* interaction; pairwise comparisons for each age group revealed that; younger participants' sketching scores were on average higher with the MixedVE than with the Abstract ($p < .01$, d = 0.21) and the Realistic VEs ($p < .001$, d = 0.34); whereas for the older participants, there were no significant differences in the sketching scores across the three visualization conditions. This finding demonstrates that the variability was too high among the older adults' performances in the VSM sample. Below we elaborate on this, and provide further interpretations of the observed results.

Discussion

In this study, we hypothesized that with a custom-designed VE (MixedVE), we can improve spatial knowledge acquisition in tasks that involve perspective switching for users with differing abilities and age groups. Specifically, we evaluated the MixedVE against two baseline alternatives (AbstractVE, RealisticVE), and examined individual and group differences on spatial knowledge acquisition as acquired from a task requiring a perspective switch with these three VEs based on age, mental rotation abilities and visuospatial memory capacity.

As hypothesized, overall, our participants were able to produce more accurate and complete sketches after having worked with the MixedVE than with the other two VEs. These results support the potential use of the MixedVE as a memory training device in navigational tasks, for spatial knowledge acquisition that involves perspective switching.

While the MixedVE improves recall accuracy, the overall task success is somewhat low, reaching around 60%. This is possibly because of the difficulty of the task to mentally switch perspectives (Taylor & Tversky, 1996), despite the help the MixedVE provides. Alternatively, given how conspicuously well-documented the difficulty of this task is, one can interpret the success levels as somewhat high: Our participants watched the videos of the routes *only once*, and still were able to draw sketches at nearly 60% accuracy and completeness on average (including their performance a week later, which brings this number down as well). Although it is not straightforward to measure how much participants might have "guessed" in a task like drawing a sketch, we interpret these results as "above chance level" (as it was theoretically proposed earlier (e.g. Montello, 1998)).

If we interpret these results as success, the fact that the learning was *intentional* may have played a role in this success (Van Asselen et al., 2006). In a real-life memory training exercise, learning would be intentional too, and repetitions would be allowed; thus the accuracy would possibly improve further.

Visuospatial memory training is relevant in all ages, but clearly more relevant as people age. It is well-understood that aging negatively correlates with success in navigational learning (e.g. Muffato et al., 2016, 2017), especially as expressed in allocentric skills (Fricke & Bock, 2018). Our findings clearly confirm the relevance of aging as a factor in spatial knowledge acquisition with perspective switching tasks: Older participants had considerably less success in accurately sketching the route they followed in any of the VEs, and this was true in both the immediate and the delayed recall stages. While aging has well-documented detrimental effects on spatial memory, one must consider that there may also be "cohort effects": that is, the younger generations are exposed to an immense technological development, and their constant use of new technology may be altering younger people's cognition in comparison to older generations (Brown, 2000). As a consequence, the so-called "Y generation" may be more prone to learning via visual, linear, and even virtual means (Schofield & Honore, 2009). Such cohort differences may have contributed to the differences in our older and younger participants' route recall scores (obtained in a technology based virtual environment). Cohort differences, however, are difficult to control. A longitudinal study might offer interesting insights, but with the fast- and constantly-changing technology, it would present other problems; comparing their learning skills with the technology relevant for example 40 years ago may not provide the pure aging effect either, because they probably would have moved on too.

Furthermore, while aging is very relevant in examining navigational memory tasks, age-related decline in visuospatial abilities and/or memory capacity can vary based on individual and group differences. On this topic, our correlation analyses (among all examined factors) with a focus on participants' scores on the MRT and VSM tests revealed interesting patterns.

The analyses based on the MRT sample indirectly suggests that the high MRT group did better than the low MRT group with the AbstractVE both in immediate and delayed recall stages (i.e. there was no significant interactions between *MRT score* x *recall stage* x *visualization*). On the one hand, it is plausible that high-MRT group would do better with AbstractVE because the MRT measures mental rotation ability in the absence of meaningful visual cues (the MRT features abstract cube drawings), and among our visualizations, the AbstractVE is the most similar to that. On the other hand, one can also take the opposite view, considering that the task heavily relied on mental rotation (perspective switch) and the MRT is designed to measure mental rotation. From this point of view, we expected to see that the high MRT group would do well with the sketching task irrespective of the visualization type. It is possible that the added (photographic) visual cues would assist the low MRT participants (thus, they "catch up" with the others), when the visual cues were present: as in a "less fit" person might do similarly well on an e-bike as fitter cyclists, because the "aid" would remove the differences). Such speculations would lend themselves well for further testing in the future.

Furthermore, the MRT analysis revealed an interaction between *age x recall stage x MRT score*. This interaction effect points to a difference in the forgetting rates: high-MRT younger participants did really well a week later (they did not seem to forget much at all), while the high-MRT older participants did not do very well after a week has passed (though note that forgetting rates did not differ for the two age groups for low-MRT participants) (Figure3(e)). Seemingly, having a high MRT relative to other older participants does not mean much for the ability to retain the acquired spatial knowledge for the older participants. However, the high-MRT older group has scores in a similar range as the younger low MRT group; for which we see a similar decline in recall in the delayed stage. This suggests that there may be an upper limit in the amount of information older people can retain based on their spatial abilities, above which they may be out of capacity. In the MRT analyses, visualization type does not interact with the other variables; that is, the relative task success with the MixedVE is constant, irrespective of age, MRT scores or recall stage.

Another interesting observation in this study is that participants' MRT and VSM scores correlate, however, the outcomes of mixed-design ANOVAs are different when we examine the MRT and VSM samples separately. We believe this difference is (at least partly) due to changes in the sample after removing the median scores in each group.

High VSM participants performed consistently better in the sketching task across all visualization types, suggesting the VSM test might be able to predict sketching performance regardless of whether photographic/visual cues exist or not on the "base maps".

Interestingly, in the VSM sample, the main effect of *age* was non-significant, implying that age-related decline in the VSM abilities is *different* to the age-related decline in the MRT abilities. Indeed, the VSM test measures at least partially the *visual* memory, and it is previously documented that visual memory might be "spared" during healthy aging (Sekuler, Kahana, McLaughlin, Golomb, & Wingfield, 2005). Furthermore, in this analysis, *recall stage x VSM score* and *visualization type x age* interacted. Recall stage x VSM score interaction suggests that the visuospatial memory capacity plays an important role in the formation of a long-term memory of the spatial configuration irrespective of visualization type. This observation contributes to our understanding of the long-term retention of spatial knowledge where perspective switching is involved. An earlier study demonstrated a gender effect (in favor of female participants) for delayed retention of survey knowledge (Witmer et al., 2002), now we see that VSM abilities of the participants may have played a role too.

The *visualization x age* interaction points to MixedVE's stronger beneficial effects for the younger participants than the older. In the case of the older participants, perhaps a "floor effect" is present, that is, overall the sketching task, especially of a route requiring multiple orientation changes (Yamamoto et al., 2018), is too difficult for them, and MixedVE's slight assistance does not suffice in this case (Moffat, 2001; Wolbers et al., 2014). This interpretation is also in line with our qualitative observations. Especially in the "delayed recall" session (~one week later), many of our older participants expressed great difficulty, and some gave up on the task.

In summary, our findings clearly confirm that spatial knowledge acquisition that involves perspective switching is a cognitively challenging task, and overall, the MixedVE makes it somewhat easier. We also have learned that as important as the visualization design is, individual and group differences *must be* considered. Based on our findings, age and visuospatial memory capacity are clearly important factors in this context. In the mid- to long-term, these observations might be useful to personalize visualizations to better fit to an individual's abilities (i.e. to create personalized (Nivala & Sarjakoski, 2005; Zipf, 2002) "memory training devices").

Conclusions and future work

In this paper, we provided new insights into the importance of addressing first of all the effect of aging, but also the individual differences in cognitive abilities (not only in terms of spatial abilities as measured by the MRT, but also of memory capacity as measured by the VSM test) in spatial knowledge acquisition in tasks that involve perspective switching when using different VEs. We gained deeper insights into the importance of visualization designs for navigational recall in general, and spatial knowledge acquisition where perspective switching is present in particular. We believe this study contributes to a better understanding of visualization design on spatial learning. Our results further help pave the way toward guidelines for designing (eventually personalized) VEs, also optimized for tasks that involve perspective switching. Such VEs can assist their users in learning a route, navigating more effectively and training their visuospatial memory both for short-term performance in spatial learning, and for long-term retention of the acquired knowledge. After these observations, our thoughts for future experimentation evolve around:

- Considering the implications of acquiring the spatial knowledge from the "reverse" perspective switch, that is encoding from a 2D view point (map reading) and decoding in a VE from a first-person perspective (i.e. wayfinding) under similar conditions as in our experiment.
- Understanding how locomotion affects learning for the different groups. Active vs. passive

involvement is known to have an effect on people (Appleyard, 1970; Chrastil & Warren, 2012). Understanding how active involvement might affect different age and ability groups can provide a more detailed reasoning as to how we should train people to achieve their best.

- How taking the route of interest more than once could affect spatial knowledge acquisition that involves perspective switching with the MixedVE as opposed to other VEs. We could then provide a benchmark as to which number of trials is the optimum for acquiring spatial knowledge.
- Decoupling age and MRT/VSM abilities. By pre-screening participants for their MRT/VSM scores, one can recruit similar numbers of people with higher and lower spatial abilities in each age group, and acquire a deeper understanding of the link between them.

Overall, based on observations we shared in this paper, we believe that individual and group differences such as age and abilities are very important to examine along with visualization design; and studies such as ours offer insights toward customized, and eventually personalized, visuospatial information displays which would facilitate route learning (as well as other learning) and potentially used for improving everyone's spatial memory.

Acknowledgments

The authors want to thank the URPP (University Research Priority Program) "Dynamics of Healthy Aging" from the University of Zurich for recruiting the older participants for the experiment. The data that support the findings of this study are available from the corresponding author, IL, upon reasonable request.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The Swiss National Science Foundation (Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung) supported this work [grant number 200021_149670] (SNSF project VISDOM).

ORCID

Ismini E. Lokka http://orcid.org/0000-0001-7970-1106
Arzu Çöltekin http://orcid.org/0000-0002-3178-3509

References

Anacta, V. J. A., Wang, J., & Schwering, A. (2014). Routes to remember: comparing verbal instructions and sketch maps. In J. Huerta, S. Schade, & C. Granell (Eds.), *Connecting a digital Europe through location and place* (pp. 311–322). Cham, Switzerland: Springer International. doi:10.1007/978-3-319-03611-3_18

Appleyard, D. (1970). Styles and methods of structuring a city. *Environment and Behavior*, *2*(1), 100–117. doi:10.1177/001391657000200106

Billinghurst, M., & Weghorst, S. (1995). The use of sketch maps to measure cognitive maps of virtual environments. *In Proceedings Virtual Reality Annual International Symposium '95* (pp. 40–47). New York: IEEE. doi:10.1109/VRAIS.1995.512478

Blades, M. (1990). The reliability of data collected from sketch maps. *Journal of Environmental Psychology*, *10*(4), 327–339. doi:10.1016/S0272-4944(05)80032-5

Brown, J. S. (2000). Growing up: Digital: How the web changes work, education, and the ways people learn. *Change: the Magazine of Higher Learning*, *32*(2), 11–20. doi:10.1080/00091380009601719

Chrastil, E. R., & Warren, W. H. (2012). Active and passive contributions to spatial learning. *Psychonomic Bulletin & Review*, *19*(1), 1–23. doi:10.3758/s13423-011-0182-x

Çöltekin, A., Francelet, R., Richter, K. F., Thoresen, J., & Fabrikant, S. I. (2018). The effects of visual realism, spatial abilities, and competition on performance in map-based route learning in men. *Cartography and Geographic Information Science*, *45*(4), 339–353. doi:10.1080/15230406.2017.1344569

Curtis, J. W. (2016). Transcribing from the mind to the map: Tracing the evolution of a concept. *Geographical Review*, *106*(3), 338–359. doi:10.1111/gere.2016.106.issue-3

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests* (Vol. 102). Princeton, NJ: Educational Testing Service.

Fricke, M., & Bock, O. (2018). Egocentric navigation is age-resistant: First direct behavioral evidence. *Current Neurobiology*, *9*(1), 69–75.

Golledge, R. G., Dougherty, V., & Scott, B. (1995). Acquiring spatial knowledge: survey versus route-based knowledge in unfamiliar environments. *Annals of the Association of American Geographers*, *85*(1), 134–158. doi: 10.1111/j.1467-8306.1995.tb01798.x

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T., & Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, *34*(2), 151–176. doi:10.1016/j.intell.2005.09.005

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*(2), 175–191. doi:10.1016/j.intell.2003.12.001

Herman, J. F., & Coyne, A. C. (1980). Mental manipulation of spatial information in young and elderly adults. *Developmental Psychology*, *16*(5), 537–538. doi:10.1037/0012-1649.16.5.537

Hurlebaus, R., Basten, K., Mallot, H. A., & Wiener, J. M. (2008). Route learning strategies in a virtual cluttered environment. In C.Freksa, N.S. Newcombe, P.

Gärdenfors, S. Wölfl (Eds.) *International Conference on Spatial Cognition* (pp. 104–120). Berlin: Springer. doi:10.1007/978-3-540-87601-4_10

Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J., & Popovich, D. L. (2015). The median split: Robust, refined, and revived. *Journal of Consumer Psychology*, *25*(4), 690–704. doi:10.1016/j.jcps.2015.06.014

Inagaki, H., Meguro, K., Shimada, M., Ishizaki, J., Okuzumi, H., & Yamadori, A. (2002). Discrepancy between mental rotation and perspective-taking abilities in normal aging assessed by Piaget's three-mountain task. *Journal of Clinical and Experimental Neuropsychology (Neuropsychology, Development and Cognition: Section A)*, *24*(1), 18–25. doi:10.1076/jcen.24.1.18.969

Ishikawa, T., & Montello, D. R. (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology*, *52*(2), 93–129. doi:10.1016/j.cogpsych.2005.08.003

Kiefer, P., Giannopoulos, I., & Raubal, M. (2014). Where am I? Investigating map matching during self-localization with mobile eye tracking in an urban environment. *Transactions in GIS*, *18*(5), 660–686. doi:10.1111/tgis.12067

Krüger, A., Aslan, I., & Zimmer, H. (2004). The effects of mobile pedestrian navigation systems on the concurrent acquisition of route and survey knowledge. In S. Brewster & M. Dunlop (Eds.) *International Conference on Mobile Human-Computer Interaction* (pp. 446–450). Berlin: Springer. doi:10.1007/978-3-540-28637-0_54

Ladd, F. C. (1970). Black youths view their environment: Neighborhood maps. *Environment and Behavior*, *2*(1), 74–99. doi:10.1177/001391657000200105

Lokka, I. E., & Çöltekin, A. (2017a). *Remembering what we see: Designing virtual environments to improve visuo-spatial recall for navigation tasks*. Proceedings of the 28th International Cartographic Conference (ICC2017), Washington, DC.

Lokka, I. E., & Çöltekin, A. (2017b). Toward optimizing the design of virtual environments for route learning: Empirically assessing the effects of changing levels of realism on memory. *International Journal of Digital Earth*, *8947*, 1–19. doi:10.1080/17538947.2017.1349842

Lokka, I. E., Çöltekin, A., Wiener, J., Fabrikant, S. I., & Röcke, C. (2018). Virtual environments as memory training devices in navigational tasks for older adults. *Scientific Reports*, *8*(1), 10809. doi:10.1038/s41598-018-29029-x

Lynch, K. (1960). *The image of the city*, Vol. 11. Cambridge, MA: MIT Press.

Meneghetti, C., Borella, E., Pastore, M., & De Beni, R. (2014). The role of spatial abilities and self-assessments in cardinal point orientation across the lifespan. *Learning and Individual Differences*, *35*, 113–121. doi:10.1016/j.lindif.2014.07.006

Meneghetti, C., Gyselinck, V., Pazzaglia, F., & De Beni, R. (2009). Individual differences in spatial text processing: High spatial ability can compensate for spatial working memory interference. *Learning and Individual Differences*, *19*(4), 577–589. doi:10.1016/j.lindif.2009.07.007

Meneghetti, C., Muffato, V., Varotto, D., & De Beni, R. (2017). How directions of route descriptions influence orientation specificity: The contribution of spatial abilities. *Psychological Research*, *81*(2), 445–461. doi:10.1007/s00426-016-0754-5

Moffat, S. (2001). Age differences in spatial memory in a virtual environment navigation task. *Neurobiology of Aging*, *22*(5), 787–796. doi:10.1016/S0197-4580(01)00251-2

Montello, D. R. (1998). A new framework for understanding the acquisition of spatial knowledge in large-scale environments. In M.J. Egenhofer & R. G. Golledge (Eds.) *Spatial and Temporal Reasoning in Geographic Information Systems*, (pp. 143–154). New York: Oxford University Press.

Moore, G. (1976). Theory and research on the development of environmental knowing. In G. T. Moore & R. G. Golledge (Eds.), *Theories, research, and methods* (pp. 138–164). Stroudsberg, PA: Dowden, Hutchinson & Ross.

Muffato, V., Della Giustina, M., Meneghetti, C., & De Beni, R. (2015). Age-related differences in pointing accuracy in familiar and unfamiliar environments. *Cognitive processing*, *16*(1), 313–317. doi:10.1007/s10339-015-0720-y

Muffato, V., Meneghetti, C., & De Beni, R. (2016). Not all is lost in older adults' route learning: The role of visuo-spatial abilities and type of task. *Journal of Environmental Psychology*, *47*, 230–241. doi:10.1016/j.jenvp.2016.07.003

Muffato, V., Meneghetti, C., Di Ruocco, V., & De Beni, R. (2017). When young and older adults learn a map: The influence of individual visuo-spatial factors. *Learning and Individual Differences*, *53*, 114–121. doi:10.1016/j.lindif.2016.12.002

Nivala, A. M., & Sarjakoski, L. T. (2005). *Adapting map symbols for mobile users*. Proceedings of International Cartographic Conference (ICC), Coruña, Spain (pp. 9–16).

O'Bryant, S. E., Humphreys, J. D., Smith, G. E., Ivnik, R. J., Graff-Radford, N. R., Petersen, R. C., & Lucas, J. A. (2008). Detecting dementia with the mini-mental state examination (MMSE) in highly educated individuals. *NIH Public Access*, *65*(7), 963–967. doi:10.1001/archneur.65.7.963.Detecting

Pazzaglia, F., & De Beni, R. (2006). Are people with high and low mental rotation abilities differently susceptible to the alignment effect? *Perception*, *35*(3), 369–383. doi:10.1068/p5465

Raubal, M., & Winter, S. (2002). Enriching wayfinding instructions with local landmarks. In M.J. Egenhofer & D.M. Mark (Eds.) *Proceedings of the Second International Conference on Geographic Information Science*, (pp. 243–259). Berlin: Springer. doi:10.1007/3-540-45799-2_17

Reichenbacher, T. (2001). Adaptive concepts for mobile cartography. *Journal of Geographical Sciences*, *11*(1), 43–53. doi:10.1007/BF02837443

Richmond, L. L., Sargent, J. Q., Flores, S., & Zacks, J. M. (2018). Age differences in spatial memory for mediated environments. *Psychology and Aging*, *33*(6), 892–903. doi:10.1037/pag0000286

Richter, K.-F., & Winter, S. (2014). *Landmarks*. Cham, Switzerland: Springer International. doi:10.1007/978-3-319-05732-3

Röser, F., Hamburger, K., Krumnack, A., & Knauff, M. (2012). The structural salience of landmarks: Results from an on-line study and a virtual environment experiment. *Journal of Spatial Science*, *57*(1), 37–50. doi:10.1080/14498596.2012.686362

Schofield, C., & Honore, S. (2009). Generation Y and learning. 360 degree. *The Ashridge Journal*, autumn, 26–32.

Sekuler, R., Kahana, M. J., McLaughlin, C., Golomb, J., & Wingfield, A. (2005). Preservation of episodic visual recognition memory in aging. *Experimental Aging Research*, *31*(1), 1–13. doi:10.1080/03610730590882800

Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology*, *43*(4), 274–310. doi:10.1006/cogp.2001.0758

Shelton, A. L., & McNamara, T. P. (2004). Orientation and perspective dependence in route and survey learning. *Journal of Experimental Psychology: Learning Memory and Cognition*, *30*(1), 158–170. doi:10.1037/0278-7393.30.1.158

Smallman, H. S., & John, M. S. (2005). Naive realism: Misplaced faith in realistic displays. *Ergonomics in Design: the Quarterly of Human Factors Applications*, *13*(3), 6–13. doi:10.1177/106480460501300303

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(1), 257–285. doi:10.1016/0364-0213(88)90023-7

Taylor, H. A., & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language*, *35*(3), 371–391. doi:10.1006/jmla.1996.0021

Thorndyke, P. W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. *Cognitive Psychology*, *14*(4), 560–589. doi:10.1016/0010-0285(82)90019-6

Van Asselen, M., Fritschy, E., & Postma, A. (2006). The influence of intentional and incidental learning on acquiring spatial knowledge during navigation. *Psychological Research*, *70*(2), 151–156. doi:10.1007/s00426-004-0199-0

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, *47*(2), 599–604. doi:10.2466/pms.1978.47.2.599

Wang, R. F. (2017). Spatial updating and common misinterpretations of spatial reference frames. *Spatial Cognition & Computation*, *17*(3), 222–249. doi:10.1080/13875868.2017.1304394

Weisberg, S. M., & Newcombe, N. S. (2016). How do (some) people make a cognitive map? Routes, places, and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(5), 768.

Wiener, J. M., Kmecova, H., & de Condappa, O. (2012). Route repetition and route retracing: Effects of cognitive aging. *Frontiers in Aging Neuroscience*, *4*, 1–7. doi:10.3389/fnagi.2012.00007

Winter, S., Raubal, M., & Nothegger, C. (2005). Focalizing measures of salience for wayfinding. In L. Meng, T. Reichenbacher, & A. Zipf (Eds.) *Map-based mobile services: Theories, methods and implementations* (pp. 125–139). Berlin: Springer. doi:10.1007/3-540-26982-7_9

Witmer, B. G., Sadowski, W. J., & Finkelstein, N. M. (2002). VE-Based training strategies for acquiring survey knowledge. *Presence: Teleoperators and Virtual Environments*, *11*(1), 1–18. doi:10.1162/105474602317343622

Wolbers, T., Dudchenko, P. A., & Wood, E. R. (2014). Spatial memory-a unique window into healthy and pathological aging. *Frontiers in Aging Neuroscience*, *6*, 35. doi:10.3389/fnagi.2014.00035

Wolbers, T., & Hegarty, M. (2010). What determines our navigational abilities? *Trends in Cognitive Sciences*, *14*(3), 138–146. doi:10.1016/j.tics.2010.01.001

Yamamoto, N., & DeGirolamo, G. J. (2012). Differential effects of aging on spatial learning through exploratory navigation and map reading. *Frontiers in Aging Neuroscience*, *4*, 1–7. doi:10.3389/fnagi.2012.00014

Yamamoto, N., Fox, M. J., Boys, E., & Ord, J. (2018). Effects of orientation change during environmental learning on age-related difference in spatial memory. *BioRxiv*, *412965*. doi:10.1101/412965

Zipf, A. (2002). User-adaptive maps for location-based services (LBS) for tourism. *Information and Communication Technologies in Tourism*, *2002*, 329–338. doi:10.1007/978-3-7091-6132-6_34

III. APPENDIX

Experimental Questionnaire

P1M

Participant No: (to be filled by the experimenter)

P1M

What was the environment you experienced?

- ○ urban
- ○ industrial
- ○ rural
- ○ I don't know

Other (please specify)

P1M

You'll be shown six pictures below. Have you seen them during your walk?

*Image 1:

1. definitely not	2. I think not	3. maybe yes/maybe not	4. I think yes	5. definitely yes	6. I don’t know
○	○	○	○	○	○

*Image 2:

1. definitely not	2. I think not	3. maybe yes/maybe not	4. I think yes	5. definitely yes	6. I don't know
○	○	○	○	○	○

*Image 3:

1. definitely not	2. I think not	3. maybe yes/maybe not	4. I think yes	5. definitely yes	6. I don't know
○	○	○	○	○	○

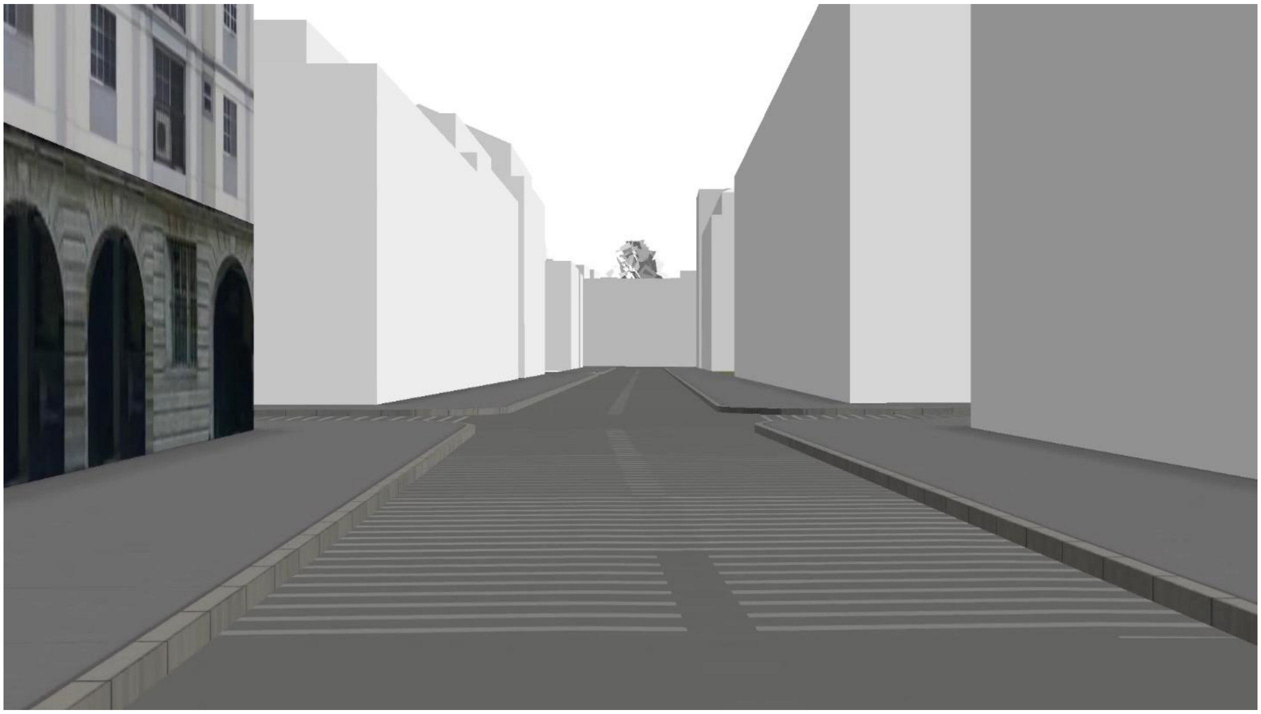


*Image 4:

1. definitely not	2. I think not	3. maybe yes/maybe not	4. I think yes	5. definitely yes	6. I don't know
○	○	○	○	○	○

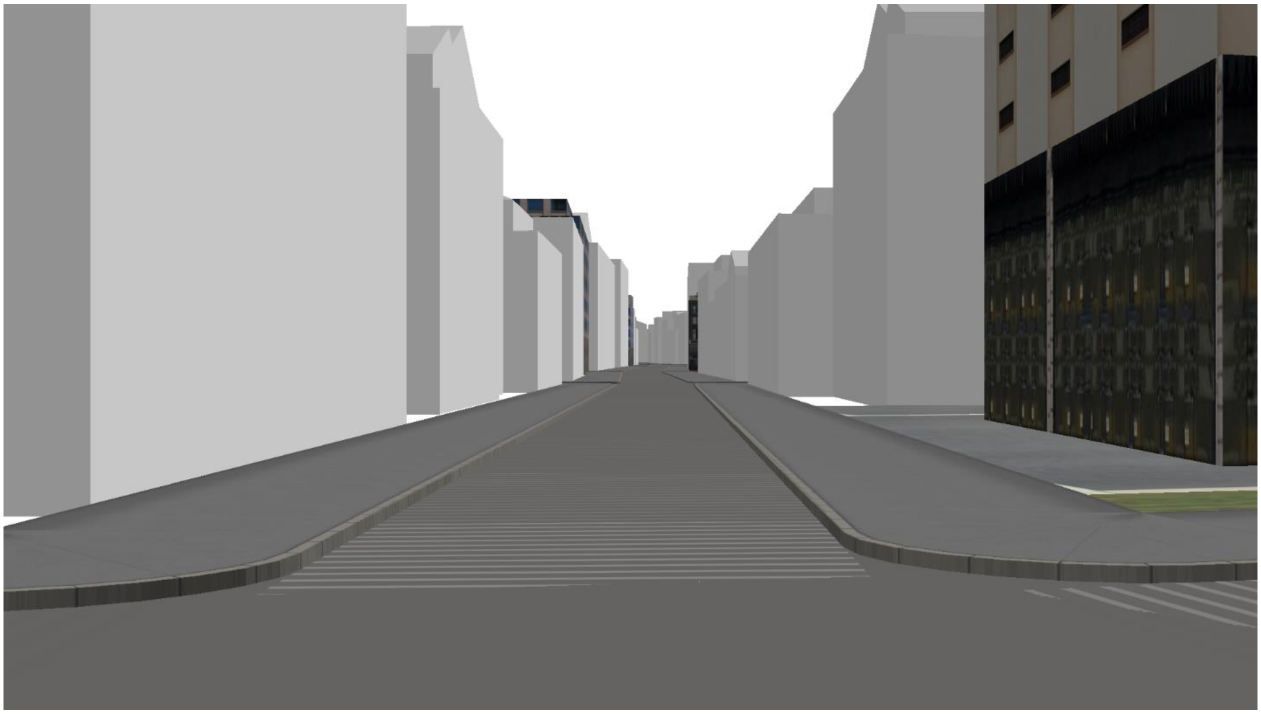


*Image 5:

1. definitely not	2. I think not	3. maybe yes/maybe not	4. I think yes	5. definitely yes	6. I don't know
○	○	○	○	○	○

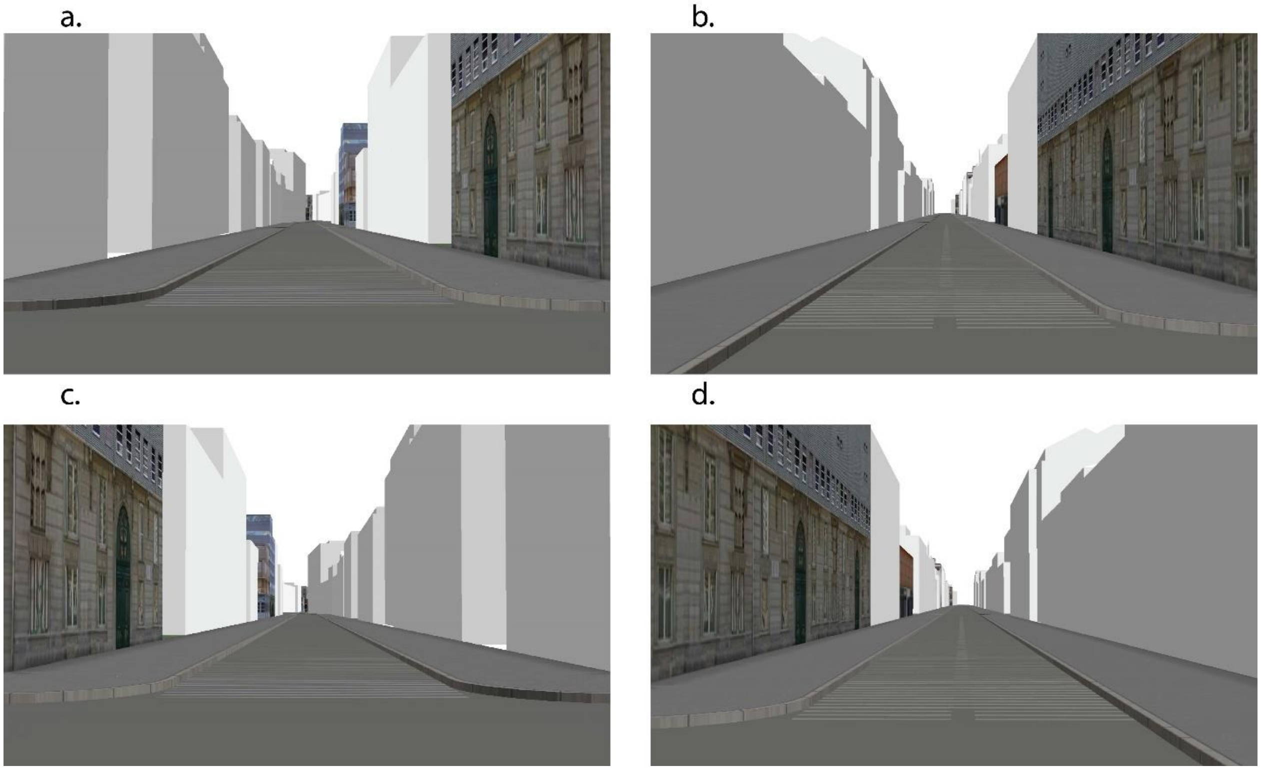
*Image 6:

1. definitely not	2. I think not	3. maybe yes/maybe not	4. I think yes	5. definitely yes	6. I don't know
○	○	○	○	○	○

P1M

* Take a look at the images below. Which of them was your starting point?

- ○ a.
- ○ b.
- ○ c.
- ○ d.
- ○ I don't know



* How <u>confident</u> are you about your answer?

- ○ 1. Not at all confident
- ○ 2.
- ○ 3. Neutral
- ○ 4.
- ○ 5. Very confident

* Take a look at the images below. Which of them was your <u>finishing point</u>?

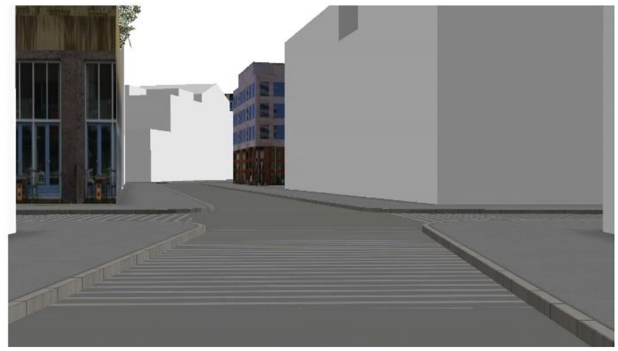
- ○ a.
- ○ b.
- ○ c.
- ○ d.
- ○ I don’t know

a.

b.

c.

d.



* How confident are you about your answer?

- ○ 1. Not at all confident
- ○ 2.
- ○ 3. Neutral
- ○ 4.
- ○ 5. Very confident

P1M

* How many times did you turn (how many turns you took) during your walk? If you don't know, please mark 0.

* How confident are you about your answer?

- ○ 1. Not at all confident
- ○ 2.
- ○ 3. Neutral
- ○ 4.
- ○ 5. Very confident

P1M

* You began your walk facing North. In the end, where were you facing?

- ○ a. North
- ○ b. South
- ○ c. West
- ○ d. East
- ○ e. I don't know

* How confident are you about your answer?

- ○ 1. Not at all confident
- ○ 2.
- ○ 3. Neutral
- ○ 4.
- ○ 5. Very confident

P1M

Please indicate the direction you followed during your walk at this intersection presented in the following seven images:

*Image 1:

a. left	b. straight	c. right	d. I don’t know
○	○	○	○

* How <u>confident</u> are you about your answer?

- ○ 1. Not at all confident
- ○ 2.
- ○ 3. Neutral
- ○ 4.
- ○ 5. Very confident

*Image 2:

a. left	b. straight	c. right	d. I don’t know
○	○	○	○

* How <u>confident</u> are you about your answer?

- ○ 1. Not at all confident
- ○ 2.
- ○ 3. Neutral
- ○ 4.
- ○ 5. Very confident

*Image 3:

a. left	b. straight	c. right	d. I don't know
○	○	○	○

* How <u>confident</u> are you about your answer?

- ○ 1. Not at all confident
- ○ 2.
- ○ 3. Neutral
- ○ 4.
- ○ 5. Very confident

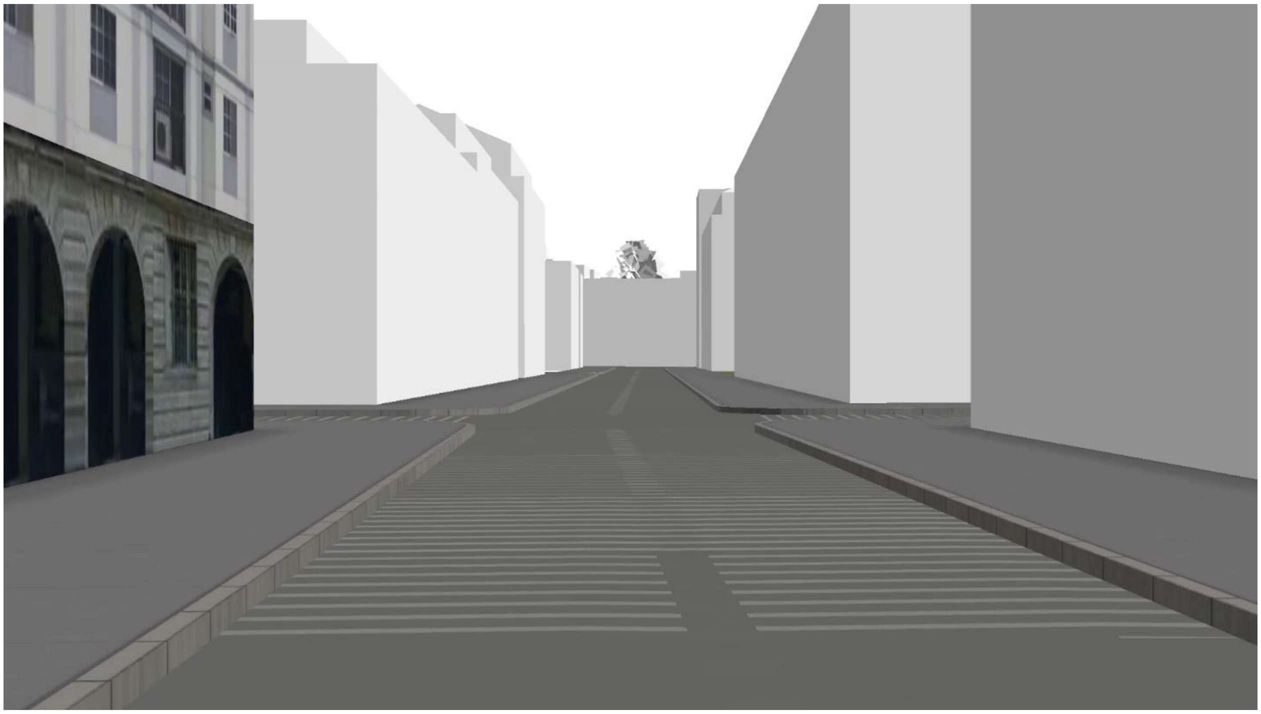


*Image 4:

a. left	b. straight	c. right	d. I don’t know
○	○	○	○

* How <u>confident</u> are you about your answer?

- ○ 1. Not at all confident
- ○ 2.
- ○ 3. Neutral
- ○ 4.
- ○ 5. Very confident

*Image 5:

a. left	b. straight	c. right	d. I don't know
○	○	○	○

* How <u>confident</u> are you about your answer?

- ○ 1. Not at all confident
- ○ 2.
- ○ 3. Neutral
- ○ 4.
- ○ 5. Very confident

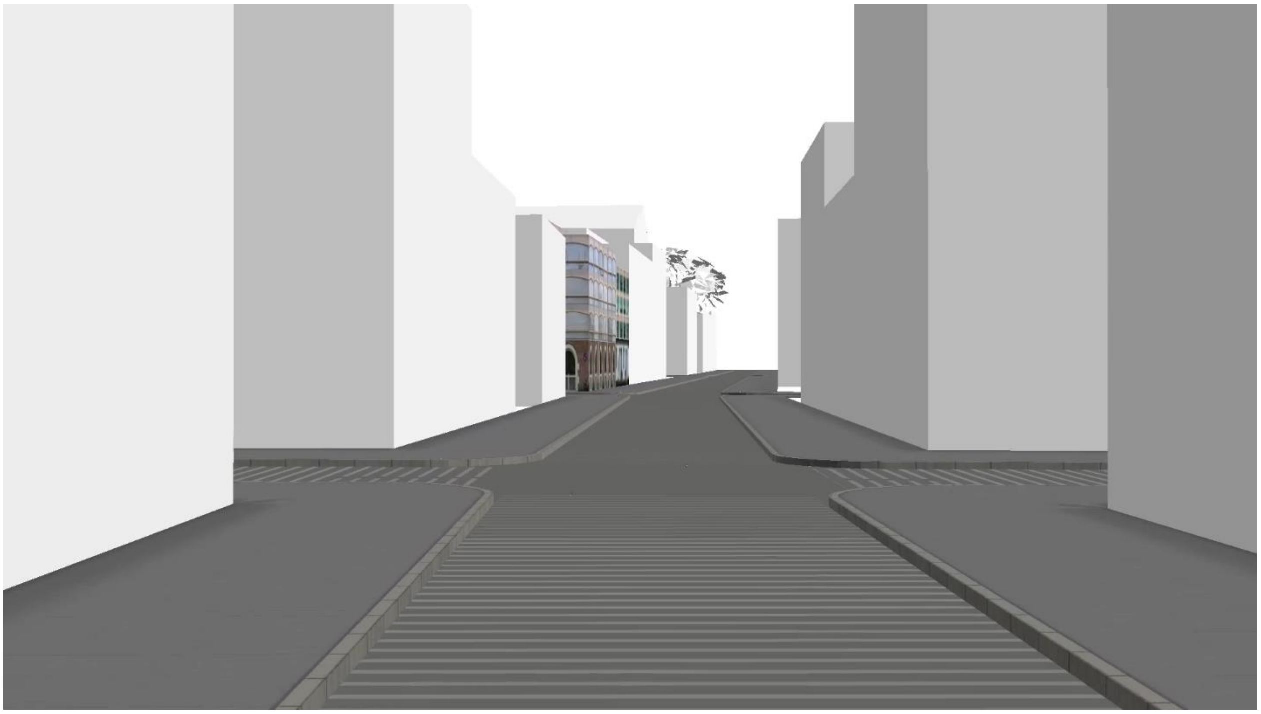


*Image 6:

a. left	b. straight	c. right	d. I don’t know
○	○	○	○

* How <u>confident</u> are you about your answer?

- ○ 1. Not at all confident
- ○ 2.
- ○ 3. Neutral
- ○ 4.
- ○ 5. Very confident

*Image 7:

a. left ○ b. straight ○ c. right ○ d. I don’t know ○

* How confident are you about your answer?

- ○ 1. Not at all confident
- ○ 2.
- ○ 3. Neutral
- ○ 4.
- ○ 5. Very confident

P1M

* Which route did you follow? Please choose one from the following four 2D views.

- ○ a.
- ○ b.
- ○ c.
- ○ d.
- ○ I don't know

* How confident are you about your answer?

- ○ 1. Not at all confident
- ○ 2.
- ○ 3. Neutral
- ○ 4.
- ○ 5. Very confident

P1M

* After watching this 3D virtual environment, overall, how <u>confident</u> do you feel about your answers to the questions about this environment?

- ○ 1. Not at all confident
- ○ 2.
- ○ 3. Neutral
- ○ 4.
- ○ 5. Very confident

Curriculum Vitae

ISMINI ELENI LOKKA

Education

2014 – 2020 Ph.D. candidate in Geographic Information Visualization and Analysis, Department of Geography, Geographic Information Science, University of Zurich (CH).

Ph.D. thesis: The impact of 3D virtual environments with different levels of realism on route learning – A focus on aging population

2007 – 2013 Diploma in Rural and Surveying Engineering, School of Rural and Surveying Engineering, National Technical University of Athens (GR).

Diploma thesis: Investigating dynamic variables with eye-movement analysis in a topographic background

Graduate courses and training

Specialization skills

- Human Computer Interaction: Cognition and usability
- Virtual Reality I
- GEO410: Geography. Matters.
- Online statistics courses
- Online R courses
- German
- Eye tracking winter school

Transferable skills

- Scientific writing
- Project Management
- Voice training & presentation skills
- Follow up on voice training and presentation skills
- Education on stage- Theater skills for presenting
- Promotion Seminar I
- Promotion Seminar II
- Principles and theory in Geography
- Resource focused stress management
- Graduate School retreat (x2)
- Managing Conflicts
- Leadership skills for doctoral candidates
- Visual communication course
- Publishing in peer-reviewed journals

Research visits & winter schools

Lokka I.E, (2015). What do we know about 3D visualizations? In *Masaryk University - Faculty of Arts. November 27th, 2015*

Lokka, I.E, & Çöltekin A. (2016). Impact of level of detail in realistic 3D geographic visualizations on memory: An empirical study with a focus on aging population using eye tracking. *In Eye Tracking – Experimental Design, Implementation, and Analysis, ETH Winter School, Monte Verita, Switzerland.* [abstract]

Published work

Peer-reviewed journal publications

Lokka I.E., Çöltekin A., (2020). Perspective switch and spatial knowledge acquisition: Effects of age, mental rotation ability and visuospatial memory capacity on route learning in virtual environments with different levels of realism. *Cartography and Geographic Information Science,* 47(1), 14-27. https://doi.org/10.1080/15230406.2019.1595151

Lokka, I. E., Çöltekin A., Wiener J., Fabrikant S.I., Röcke C. (2018) Virtual environments as memory training devices in navigational tasks for older adults. *Scientific reports* 8(1). https://doi.org/10.1038/s41598-018-29029-x

Lokka I.E., Çöltekin A., (2019). Towards optimizing the design of virtual environments for route learning: empirically assessing the effects of changing levels of realism on memory. *International Journal of Digital Earth*, 12(2), 137-155. https://doi.org/10.1080/17538947.2017.1349842

Roth R.E., Çöltekin A., Delazari L.S., Filho H.F., Griffin A., Hall A., Korpi J., **Lokka I.E.**, Mendonça A.L.A., Ooms K., van Elzakker K.P.J.M., (2017). User Studies in Cartography: Opportunities for Empirical Research on Interactive Maps and Visualizations. *International Journal of Cartography,* 3(1), 61-89*.* https://doi.org/10.1080/23729333.2017.1288534

Journal peer-reviewer

- **Scientific Reports – Nature** https://www.nature.com/srep/
- **International Journal of Geographic Information Science** https://www.tandfonline.com/toc/tgis20/current

Conference publications

Lokka I.E., Çöltekin A., (2019). Age differences in attention and memory in a virtual reality route learning task, In *the 5th International Conference Aging & Cognition 2019, Zurich, Switzerland (April 2019).* [abstract]

Lokka, I. E., & Çöltekin, A. (2018). Evaluating route learning performance of older and younger adults in differently-designed virtual environments: A task-differential analysis. In *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Delft. Netherlands*. October 1st-5th, 2018.

Lokka, I. E., & Çöltekin, A. (2018). Do age differences affect performance in 2D sketching based on a first-person perspective (3D) route learning task in differently-designed virtual environments? In *Spatial Cognition 2018. Tuebingen, Germany*, September 05-08, 2018.

Lokka I.E., Çöltekin A., (2018). Virtual environments designed to improve route learning performance: A focus on age and visuospatial abilities. In *ICA Commissions Joint Workshop Atlases, Cognition, Usability*, *April 2018.*

Lokka I.E., Çöltekin A., (2018). A virtual reality experiment for improving the navigational recall: What can we learn from eye movements of high- and low-performing individuals*.* In *Proceedings of the 3rd International Workshop on Eye Tracking for Spatial Research (January 2018).*

Lokka I.E., Çöltekin A., (2017). Remembering what we see: Designing virtual environments to improve visuo-spatial recall for navigation tasks. In *Proceedings of the 28th International Cartographic Conference (ICC2017), Washington DC, USA (July, 2017).* [short paper]

Roth R.E., Çöltekin A., Delazari L.S., Filho H.F., Griffin A., Hall A., Korpi J., **Lokka I.E.**, A.L. Mendonça, Ooms K., van Elzakker C.P.J.M., (2017). User Studies in Cartography: A Collaborative Research Agenda. In *Proceedings of the 27th International Cartographic Conference (ICC2017), Washington DC, USA (July, 2017).* [short paper]

Lokka, I.E., Çöltekin, A. (2017). Navigational learning in virtual environments that are designed to improve memory– Individual and group differences based on spatial abilities and age. Workshop on Urban Wayfinding and the Brain, June 14th, 2017, UCL, London, UK. [abstract, poster presentation]

Lokka, I.E., Çöltekin, A., (2017). Designing memorable 3D geovisualizations for the elderly. *In the 4th International Conference Aging & Cognition 2017, Zurich, Switzerland (April 2017).* [abstract]

Lokka, I.E., Çöltekin, A., (2017). Virtual environments as memory training devices for navigational tasks as we age: A design perspective. In *Second International Workshop on Models and Representations in Spatial Cognition, Tübingen*. [abstract]

Lokka, I.E., Çöltekin, A. (2016). Simulating navigation with virtual 3D geovisualizations–A focus on memory related factors. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. [short paper]

Çöltekin A., **Lokka I.E.,** & Zahner M. (2016). On the usability and usefulness of 3D (geo)visualizations—A focus on virtual reality environments. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. [short paper]

Çöltekin A., **Lokka I.E.,** & Boér A. (2015). The Utilization of Publicly Available Map Types by Non-Experts–A Choice Experiment. In *Proc. of the 27th International Cartographic Conference, Rio de Janeiro.* [full paper]