Smart Map Augmented: Exploring and Learning Maritime Audio-Tactile Maps without Vision: The Issue of Finger or Marker Tracking

Mathieu Simonnet

Article

Smart Map Augmented: Exploring and Learning Maritime Audio-Tactile Maps without Vision: The Issue of Finger or Marker Tracking

Mathieu Simonnet

IMT Atlantique, LEGO, 29200 Brest, France; mathieu.simonnet@imt-atlantique.fr

Abstract: Background: When exploring audio-tactile nautical charts without vision, users could trigger vocal announcements of a seamark's name thanks to video tracking. In a first condition they could simply use a green sticker fastened at the tip of a finger and in a second condition they could handle a small handy green object, called the marker. Methods: In this study, we attempted to compare finger and marker tracking conditions to complete spatial tasks without vision. More precisely, we aimed to better understand which kind of interaction was the most efficient to perform either localization or estimation of distance and direction tasks. Twelve blindfolded participants realized these two spatial tasks on a 3D-printed audio-tactile nautical chart. Results: Results of the localization tasks revealed that in finger condition, participants were faster in finding geographic elements, i.e., seamarks. During estimation tasks, no differences were found between the precision of distances and direction estimations in both conditions. However, spatial reasoning took significantly less time in marker condition. Finally, we discussed the issue of the efficiency of these two interaction conditions depending on the spatial tasks. Conclusions: More experimentation and discussion should be undertaken to identify better modalities for helping visually impaired persons to explore audio-tactile maps and to prepare navigation.

1. Introduction

1.1. Context

The use of new technologies to improve the accessibility of geographic maps for visually impaired persons (VIP) has been a major challenge of the past 20 years [1,2]. From audio-tactile tablets [3] to virtual reality feedback [4], including 3D printed maps [5] and tangible interactions [6], many studies have provided VIPs with more and more effective systems, since they have increasingly involved users in the design and evaluation stages [7]. This is the case of the SMAug (Smart Map Augmented) project, whose specifications for 3D printed audio-tactile nautical charts were produced in cooperation with a blind sailor.

1.2. Objective

The objective of the SMAug project is to offer collaborative nautical charts to help everyone, including VIPs, to discover and memorize the maritime environment. As the most important geographical elements at sea are navigation aids, such as buoys, beacons and lighthouses, tactile nautical charts must follow the IHO (International Hydrographic Organization) (https://iho.int/ accessed on 6 June 2022) visual design to make maritime buoyage recognizable by sighted sailors. Then, the ten types of seamarks should also be understandable by touch. In this work, we used the results related to 3D-printed tactile nautical charts from a previous study, [8] and added audio information thanks to a visual tracking system. More precisely, users digitally explored 3D maps, a camera was mounted above the tactile map, and a colorimetric-based detection script allowed voice

announcements to be triggered to give the names of the seamarks. Thus, the accessibility of a tangible tactile map with clear shapes and the power of an internet database with massive geo-referenced information were combined. The SMAug system was made to be used at a chart table aboard a sailboat to prepare navigation, but it can also be used at home or in the office (cf. Figure 1).
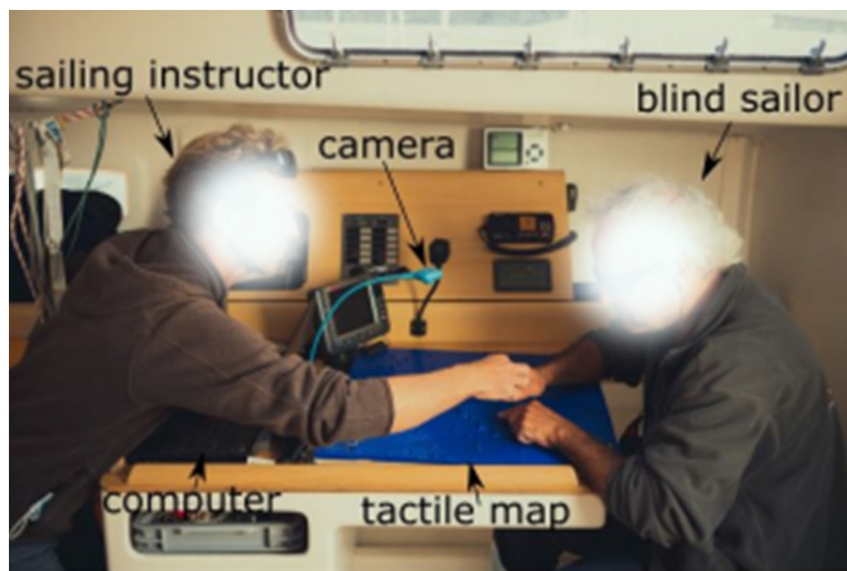

Figure 1. SMAug system at a chart table on a sailboat before navigation. The blind sailor consults the tactile map with a sticker on a finger to track it. Vocal announcements are triggered by the SMAug software. Thus, the sailing instructor and the blind sailor could prepare navigation together on the same map.

1.3. Spatial Cognition and Tactile Exploration

Considering previous knowledge on spatial cognition [9], to understand the spatial layout of a map and/or an environment, we first need to perceive different landmarks, or seamarks, as unique and salient elements. Secondly, we need to identify the relationships between the landmarks. This process is required to mentally encode and remember spatial representations, the so-called spatial cognitive map [10]. Without vision, VIPs sequentially discover the elements and move their hands to explore and estimate the distance and direction between them. Therefore, memorizing global spatial layouts requires that VIPs widely use the haptic (i.e., tactile-kinesthetic) modality to recognize elements and remember the spatial relationships at the same time [11].

The lack of vision leads to the encoding of these spatial relationships in an egocentric spatial frame of reference (i.e., body-to-objects). However, with a more important cognitive load, allocentric representation (i.e., object-to-object) is also accessible using the haptic modality [12]. This is important, since the use of maps requires the coordination of both spatial frames of reference [13]. Moreover, VIPs also tend to manage categorical (left/right) spatial relations more easily than coordinate (metric) spatial relations, although the latter are more precise [14,15]. Thus, assessing allocentric and coordinate spatial representation is particularly interesting.

During manual exploration, "tactile fixations", like eye fixations, have been observed. More precisely, when a finger stops on an element of a map, this draws specific attention to this element and its spatial relationships [16]. Therefore, one could surmise that vocal information should help VIPs to explore a spatial layout. This requires appropriate names and expected spatial relationships. Vocal information should also be triggered at the right moment, i.e., when attention is available to remember names or relationships required to build and/or complete a mental model.

1.4. Interaction Design

To help visually impaired persons to identify different elements as unique and salient landmarks, we made seamarks audible. Their names were vocally announced when proximity was detected. As mentioned above, we followed the colored marker using video tracking algorithms (cf. Figure 2). The originality of this work was that the marker can be a sticker fastened to the tip of the finger, as previously tested [6], or a handy manipulable piece that users could move next to the tactile elements to launch audio information. Both finger and marker conditions aim to give to VIPs the expected information using the same video tracking algorithm. However, focusing on the users' behaviors, the finger condition (i.e., green sticker) naturally launches vocal information during haptic exploration, whereas the marker condition (i.e., green piece) only triggers a vocal announcement once the user has intentionally moved the manipulable object toward the place he/she wants to get information about.
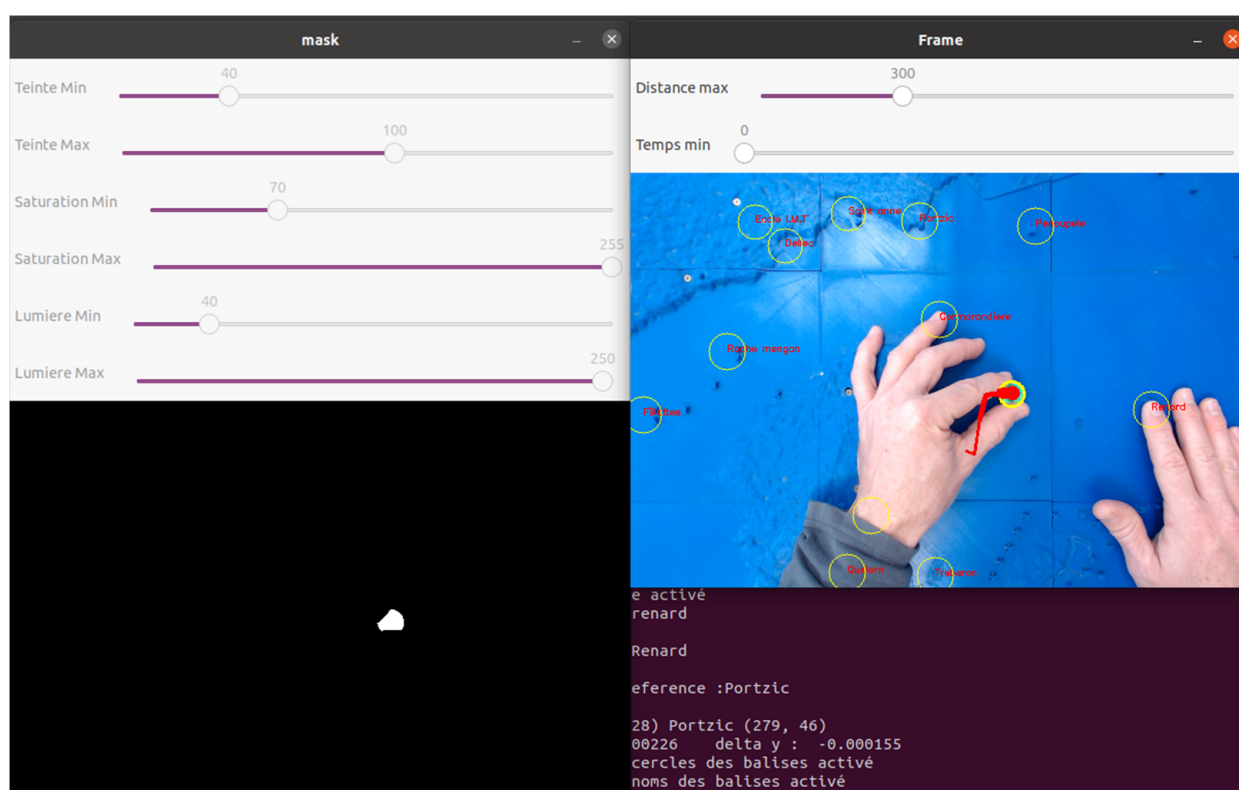

Figure 2. On the left, the mask window shows the binarized image that filters color using the HSV parameter, current settings corresponding to the green color. On the right, the frame windows show the RGB image. The 12 thin circles with names in small print represent seamark areas, and the wide circle in the middle of the finger indicates the tracked color. When the marker enters a seamark area, the corresponding name is vocally announced.

Thus, the SMAug system currently offers two different conditions of interaction to query information on seamarks, and we expect to understand the benefits and inconveniences of both.

1.5. Research Question

A preliminary question therefore arises: During tactile map exploration, were vocal announcements more efficient when tracking the tip of the finger or when tracking the manipulable marker? More precisely, which type of spatial task took better advantage of which condition?

Taking into account previous works, we hypothesize that the marker condition improves allocentric and coordinate spatial representation. Indeed, positioning the green piece at different seamark locations should require memorizing the seamarks' positions more thoroughly, and then fostering object-to-object representation.

2. Materials and Methods

2.1. Participants

Twelve blindfolded engineering students took part in this experiment. Six females and six males were blindfolded during the tasks. Participants were between 22 and 24 years old (mean = 22.5). All participants were familiar with the technologies tested, and were right-handed.

In response to the question, "Are you familiar with paper charts?", six answered "medium", four, "no", and two, "yes".

In response to the question, "Are you familiar with maps on screen?", ten said "yes" and two, "medium".

In response to the question, "Are you familiar with numerical cardinal direction?", six answered "medium", four, "no", and two, "yes".

2.2. Equipment

Participants explored the same 60 × 60 cm 3D-printed tactile map. It was a real maritime chart for blind sailors. It represented the bay of Brest (France) at a scale of 1:20,000, i.e., 1 cm on the map represented 200 m in the real environment. Land elevation followed the altitude of the digital terrain model, but all the values were doubled to make the coastline easier to feel by touch. The sea area was flat and contained 38 seamarks among 10 kinds of marks. The experiment used only 12 seamarks from 4 kinds of shape. In other words, only 12 seamarks were made audible. The map model was made with the open-source software Navisu (https://github.com/terre-virtuelle/navisu accessed on 6 June 2022) Nine tiles of 20 × 20 cm were printed with an Ultimaker 3 3D-printer with blue PLA material.

To run the colorimetric-based detection SMAug software, we used a dual core computer with 4 GB RAM, the Ubuntu exploitation system, and a 1080 HD Logitech camera mounted on a 50 cm-high desk lamp. The software application was coded with Python 3, and the libraries OpenCV (https://opencv.org/ accessed on 6 June 2022) and pyttsx3 (https://pypi.org/project/pyttsx3/ accessed on 6 June 2022) for computer vision and text-to-speech. Vocal announcements were made via loudspeaker.

The finger condition used a 15-mm-diameter sticker on the right index finger, and the marker condition used a 17-mm-diameter 3D-printed hand piece (cf. Figures 3 and 4).

Figure 3. In the finger condition, a 15 mm green sticker is fastened at the tip of the right index finger.
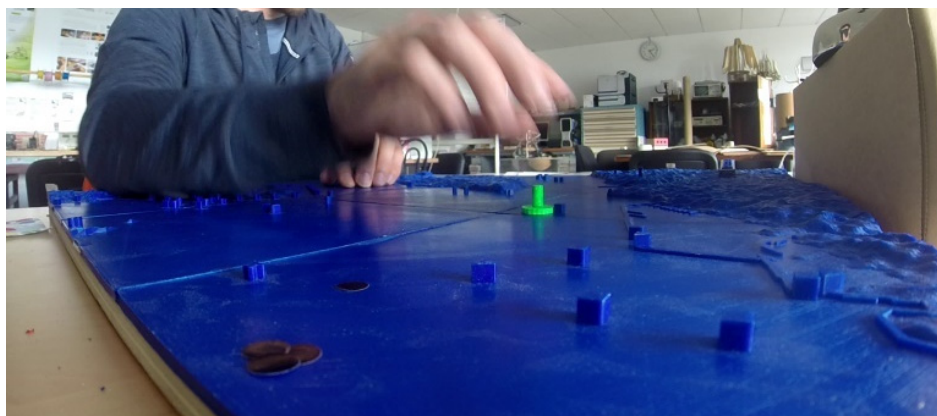
Figure 4. In the marker condition, the green 17 mm movable object had to be manipulated by the participants.

In each condition, as soon as the finger or the marker was less than 1 cm from the seamark, the vocal announcement of the seamark's name was triggered once. To repeat the vocal announcement, users had to move the right index finger or the marker further than 1 cm away, and then back in again.

3. Experimental Tasks and Data Collection

The 12 participants first performed a training run to discover the 12 seamarks. This training was done twice: once in the finger condition, and once in the marker condition. During this sequence, we did not collect any data.

During the subsequent localization and estimation tasks, the participants' explorations were totally free. No instructions indicated how to manually explore the layout or imposed restrictions on how to handle the marker.

3.1. Localization Task

Thereupon each participant performed the localization task. They had to find the seamark indicated by name. Three seamarks in the finger condition, three in the marker condition; then three in the finger condition again, and three more in the marker condition. For example, the question was: "Could you find Penoupele?" Here we collected data on how long it took to find each seamark. Naturally, conditions were counterbalanced to avoid the order effect.

3.2. Estimation Task

The next task was to estimate the distance and the direction from one seamark to another. Three in the finger condition, three in the marker condition, and so on. For example, the question could be, "What is the distance and direction between Penoupele and Portzic?" Distance estimations were given in meters, corresponding to the scale, and direction estimations in cardinal degrees (i.e., 0° north, 90° east, 180° south, 270° west). Users could use the SMAug system during the question. Thus, they did not have to remember the names of the seamarks. We collected the precision errors of the estimations of distance and direction. Finally, we also collected data on how long it took to find the seamarks, and then to compute the calculation and give the answer.

At the end, we ask the users to complete a short questionnaire on which conditions they preferred for each task: localization and estimation.

4. Results

Since the data we collected did not follow the normal distribution, we performed nonparametric statistical tests. More precisely, we compared response times, distance errors, and direction errors using the Wilcoxon test. As the use of the 12 seamarks was

also counterbalanced between conditions to avoid the configuration effect, we applied the non-pairwise Wilcoxon test.

The Figures 5–8 display the results within boxplots. Here, the grey box shows the interquartile range (IQR). The bold line inside represents the median. The thin line at the bottom is the minimum. The thin line at the top is the maximum. The small circles are the outliers. One star between two boxplots indicates a significant difference ($p < 0.05$) and two stars mean that the difference is very significant ($p < 0.01$).

4.1. Localization Task

When participants had to locate the six seamarks in each condition, they spent less time in the finger condition (median = 10 s) than in the marker condition (median = 17.5 s). This difference was very significant (W = 1517.5, $p < 0.01$) (cf. Figure 5). As a matter of fact, this result showed that the tracking finger technique allowed the blindfolded users to locate different elements on tactile maps faster than the object tracking technique.
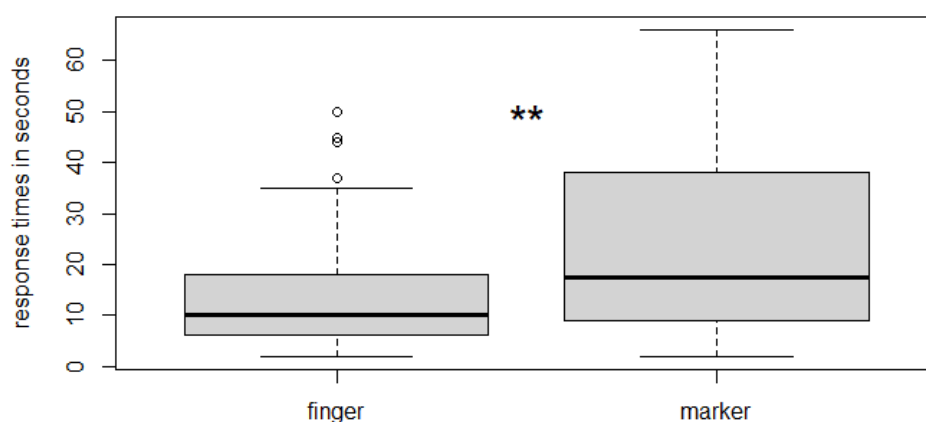

Figure 5. Response times in the finger and marker conditions.

4.2. Estimation Task

4.2.1. Distance

When participants had to estimate the distances between the seamarks, the distance estimation errors were slightly higher (median = 600 m) in the finger condition than in the marker condition. However, no significant difference was found (W = 2742.5, $p > 0.05$) (cf. Figure 6).
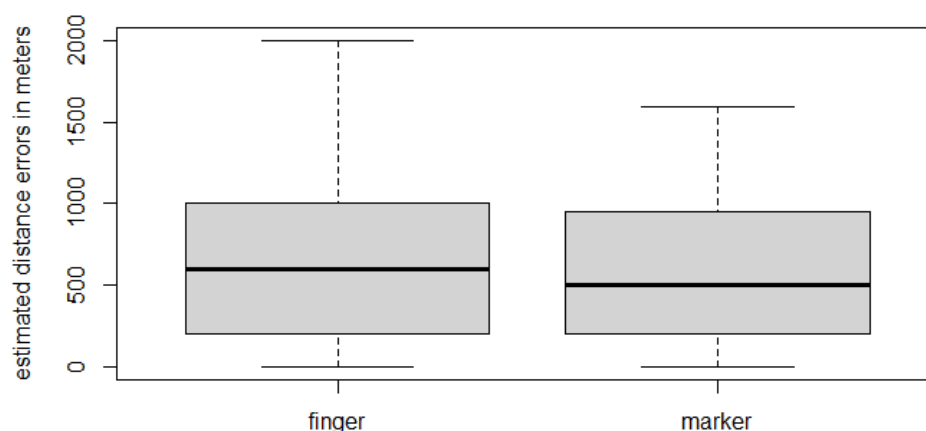

Figure 6. Estimated distance errors between seamarks in the finger and marker conditions.

4.2.2. Directions

Similarly, when participants had to estimate the directions between the seamarks, no significant differences were found ($W = 2645.5$, $p > 0.05$) (cf. Figure 7). We even noticed than the estimated directions errors were the same in both conditions (median = 9°).
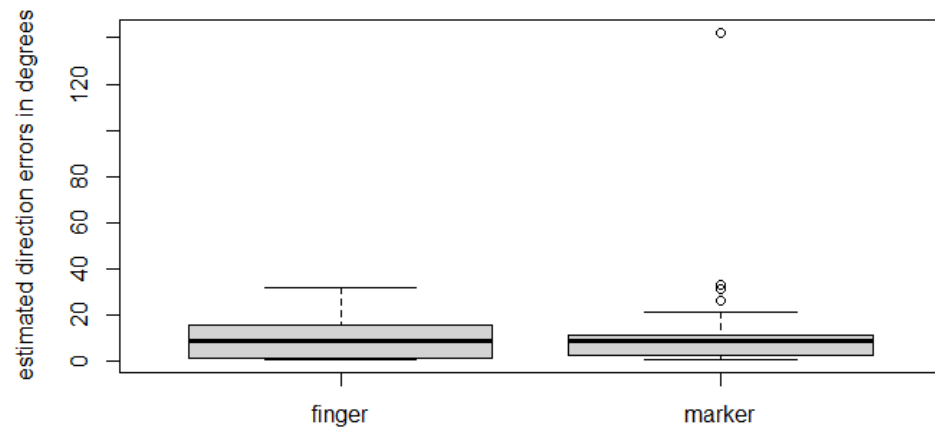

Figure 7. Estimated direction errors between seamarks in the finger and marker conditions.

Thus, the finger or marker conditions did not seem to have an impact on the distance and direction estimations.

4.2.3. Response Times during the Estimation Task

When participants were asked to estimate distances and directions between seamarks, they first found the elements on the 3D-printed tactile map, and then thought deeply about the answers they would give to the experimenter to estimate distance and direction. We recorded how long it took to find the seamarks and to give an answer.

During this estimation task, participants spent only slightly less time locating the two seamarks in the finger condition (median = 18 s) than in the marker condition (median = 21 s). Contrary to the first results in the localization task, no significant difference appeared ($W = 750$, $p > 0.05$).

Then, the time spent to estimate the distance and direction was significantly longer ($W = 1210.5$, $p < 0.05$) in the finger condition (median = 45.5 s) than in marker condition (median = 35.5 s) (cf. Figure 8).

Thus, the marker condition seemed faster than the finger condition for estimating distance and direction.
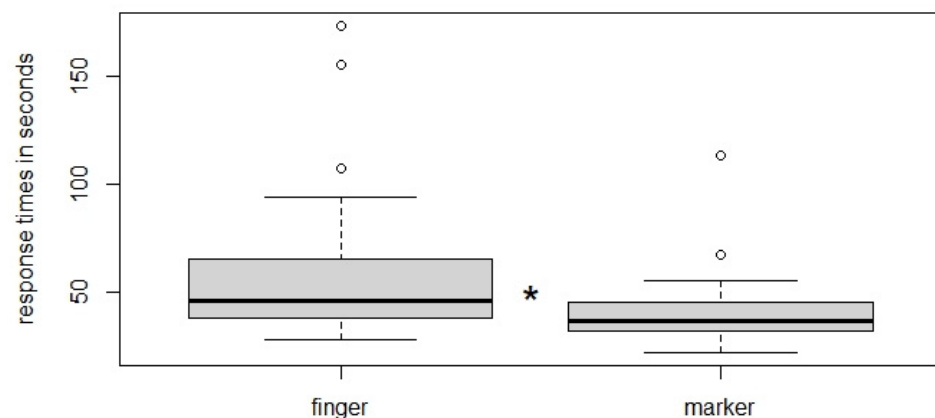

Figure 8. Time required to estimate the distance and direction between seamarks after having found them.

5. Discussion

The question was to identify which exploration condition, the finger or marker, was the most efficient, i.e., faster and more precise, for locating geographical elements and perceiving the distance and direction between them in an allocentric and coordinate spatial representation. The two interaction conditions proposed two slightly different kinds of exploration. Indeed, the tracked finger condition led to listening to the marked index. It enabled the tactile fixation concept revealed by Zhao et al. [12]. In contrast to the "talking finger", the use of the marker, i.e., tracked manipulable object, naturally encouraged exploring the layout while both hands are free, and then finding the object only when vocal information was expected. Indeed, every participant spontaneously used the strategy of putting the marker on the side of the tactile map at the same place, to avoid losing too much time in finding it again.

The first result of the localization task revealed that participants were faster to locate the six designated seamarks in the finger condition. This could be explained by the time required to pick up and bring the marker close to the focused seamark. This result was expected and mainly confirmed that the detection system has worked properly. This also emphasized the choice of Shi et al. [6] in the Molder system. Indeed, they gave effective access to the tactile map info to VIPs using the same finger condition. In the case of SMAug, which consisted of exploring nautical charts without vision, it is particularly interesting to use this finger condition when searching for the seamark's names on board at the chart table, before or during sailing. Although some participants sometimes said "OK, OK, I know ... " when the name of the seamark was automatically announced, they did not claim to be disturbed by the automatic vocal announcement when we questioned them about this first task.

The second result showed that participants did not take any advantage of the marker condition to estimate the distance and direction between the different seamarks. Indeed, the distance and direction errors were similar in both conditions. This means that the marker condition did not improve allocentric and coordinate spatial representation as we hypothesized. Thus, positioning the green piece at different seamark places did not seem to require memorizing the seamarks' positions more thoroughly and improving object-to-object representation. This could be explained by the simplicity of the spatial task and/or the layout. When we observed participants' explorations, they only focused on the two seamarks involved in the estimation task. Sometimes they just remembered the names and did not even use the marker, and at other times they moved the marker to check whether the seamark was the right one. In any case, they always performed both hands' exploration where they touched the two seamarks and attempted to measure distance and direction with the fingers, hands, and even arms depending on the participants. No different exploratory patterns were identified depending on the condition. Here we could corroborate the observations made by Zhao et al. [16]. since finger on the seamarks were long even when participants were thinking about something else like scale conversion or cardinal orientation cogitation. Haptic exploratory patterns should be more deeply investigated to better understand the relation between spatial tasks, interaction types (i.e., conditions), and underlying cognitive processes.

The last result showed that once having found the two seamarks, the participants spent less time estimateing distance and direction in the marker condition than in the finger condition. This could be explained by the intermittent repetition of the name of the seamark under the tracked finger when touching the seamark in question, or when measuring with fingers to give an estimation. Indeed, we observed some participants being stopped in their calculation because of an untimely vocal announcement. Moreover, some participants explained that the automatic announcement could be disturbing for this task. This suggested that as soon as the spatial task requires deeper reasoning, it could be interesting to propose the vocal announcement only on demand. Otherwise, tactile fixation intention should be analyzed more precisely to avoid participants' intention misinterpretation. Another good point for the marker condition was to observe the participants looking

for the movable piece, and then coming back to the tactile map. Thus, as we mentioned in the research question subsection, the marker could have the potential to foster allocentric and coordinate spatial representation. However, at the moment, one could only suggest that this back-and-forth required a deeper cognitive effort to find the seamark and therefore may stimulate spatial layout learning. This would corroborate the findings of Waller who showed that one manner to improve spatial learning of a configuration was to get lost first [17]. However, any conclusion about these two conditions would require performing more experiments with more participants and more questions including VIPs, and maybe with more complex configurations and tasks in the ego and allocentric spatial frame of references. Thus, these first results could pave the way for promising future work.

References

1. Brock, A.; Truillet, P.; Oriola, B.; Jouffrais, C. Usage of Multimodal Maps for Blind People: Why and How. In *ACM International Conference on Interactive Tabletops and Surfaces—ITS '10*; ACM Press: New York, NY, USA, 2010; p. 247. [CrossRef]
2. Froehlich, J.E.; Brock, A.M.; Caspi, A.; Guerreiro, J.; Hara, K.; Kirkham, R.; Schöning, J.; Tannert, B. Grand Challenges in Accessible Maps. *Interactions* **2019**, *26*, 78–81. [CrossRef]
3. Goncu, C.; Marriott, K. GraVVITAS: Generic Multi-Touch Presentation of Accessible Graphics. *Lect. Notes Comput. Sci.* **2011**, *6946 LNCS*, 30–48. [CrossRef]
4. Simonnet, M.; Jacobson, D.; Vieilledent, S.; Tisseau, J. SeaTouch: A Haptic and Auditory Maritime Environment for Non Visual Cognitive Mapping of Blind Sailors. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2009; Volume 5756 LNCS, pp. 212–226. [CrossRef]
5. Holloway, L.; Marriott, K.; Butler, M. Accessible Maps for the Blind: Comparing 3D Printed Models with Tactile Graphics. In Proceedings of the Conference on Human Factors in Computing Systems Proceedings, Honolulu, HI, USA, 25–30 April 2020; Association for Computing Machinery: New York, NY, USA, 2018; pp. 1–13. [CrossRef]
6. Shi, L.; Zhao, Y.; Gonzalez Penuela, R.; Kupferstein, E.; Azenkot, S. Molder: An Accessible Design Tool for Tactile Maps Visual Impairments; Tactile Maps; Design Tool CSS Concepts. In *Proceedings of the Conference on Human Factors in Computing Systems Proceedings, Honolulu, HI, USA, 25–30 April 2020*; Association for Computing Machinery: New York, NY, USA, 2020. [CrossRef]
7. Albouys-Perrois, J.; Laviole, J.; Briant, C.; Brock, A.M. Towards a Multisensory Augmented Reality Map for Blind and Low Vision People: A Participatory Design Approach. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; Association for Computing Machinery: New York, NY, USA, 2018. [CrossRef]
8. Simonnet, M.; Morvan, S.; Marques, D.; Ducruix, O.; Grancher, A.; Kerouedan, S. Maritime Buoyage on 3D-Printed Tactile Maps. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, Galway, Ireland, 22–24 October 2018; pp. 450–452. [CrossRef]
9. Siegel, A.W.; White, S.H. The Development of Spatial Representations of Large-Scale Environments. In *Advances in Child Development and Behavior*; Reese, H., Ed.; Academic Press: New York, NY, USA, 1975; pp. 10–55.
10. Kitchin, R.M. Cognitive Maps: What Are They and Why Study Them? *J. Environ. Psychol.* **1994**, *14*, 1–19. [CrossRef]
11. Hatwell, Y.; Streri, A.; Gentaz, E. *Touching for Knowing: Cognitive Psychology of Haptic Manual Perception*; John Benjamins Publishing: Amsterdam, The Netherlands, 2003.
12. Millar, S. *Understanding and Representing Space: Theory and Evidence from Studies with Blind and Sighted Children*; Oxford University Press: Oxford, UK, 1994.
13. Gaunet, F.; Thinus-Blanc, C. Early-Blind Subjects' Spatial Abilities in the Locomotor Space: Exploratory Strategies and Reaction-to-Change Performance. *Perception* **1996**, *25*, 967–981. [CrossRef] [PubMed]
14. Kosslyn, S.M. *Image and Brain: The Resolution of the Imagery Debate*; The MIT Press: Cambridge, MA, USA, 1994.
15. Ruggiero, G.; Ruotolo, F.; Iachini, T. Egocentric/Allocentric and Coordinate/Categorical Haptic Encoding in Blind People. *Cogn. Process.* **2012**, *13* (Suppl. 1), 313–317. [CrossRef]

16. Zhao, K.; Bardot, S.; Serrano, M.; Simonnet, M.; Oriola, B.; Jouffrais, C. Tactile Fixations: A Behavioral Marker on How People with Visual Impairments Explore Raised-Line Graphics. In Proceedings of the CHI'21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; p. 12. [CrossRef]
17. Waller, D.; Hodgson, E. Transient and Enduring Spatial Representations under Disorientation and Self-Rotation. *J. Exp. Psychol. Learn. Mem. Cogn.* **2006**, *32*, 867–882. [CrossRef] [PubMed]