Article

Cartographic Visualization for Indoor Semantic Wayfinding

Nikolaos Bakogiannis *, Charalampos Gkonos and Lorenz Hurni

Institute of Cartography and Geoinformation, ETH Zurich, 8093 Zurich, Switzerland;
gkonosch@gmail.com (C.G.); lhurni@ethz.ch (L.H.)
* Correspondence: n.bakogiannis@outlook.com

check for
updates

Abstract: In recent years, pedestrian navigation assistance has been used by an increasing number of people to support wayfinding tasks. Especially in unfamiliar and complex indoor environments such as universities and hospitals, the importance of an effective navigation assistance becomes apparent. This paper investigates the feasibility of the indoor landmark navigation model (ILNM), a method for generating landmark-based routing instructions, by combining it with indoor route maps and conducting a wayfinding experiment with human participants. Within this context, three different cartographic visualization scenarios were designed and evaluated. Two of these scenarios were based on the implementation of the ILNM algorithm, with the concurrent effort to overcome the challenge of representing the semantic navigation instructions in two different ways. In the first scenario, the selected landmarks were visualized as pictograms, while in the second scenario, an axonometric-based design philosophy for the depiction of landmarks was followed. The third scenario was based on the benchmark approach (metric-based routing instructions) for conveying routing instructions to the users. The experiment showed that the implementation of the ILNM was feasible, and, more importantly, it was beneficial in terms of participants' navigation performance during the wayfinding experiment, compared to the metric-based instructions scenario (benchmark for indoor navigation). Valuable results were also obtained, concerning the most suitable cartographic approach for visualizing the selected landmarks, while implementing this specific algorithm (ILNM). Finally, our findings confirm that the existence of landmarks, not only within the routing instructions, but also as cartographic representations on the route map itself, can significantly help users to position themselves correctly within an unfamiliar environment and to improve their navigation performance.

Keywords: indoor navigation assistance; landmarks in routing instructions; indoor landmark navigation model; pictograms

1. Introduction

Pedestrian navigation aids are used to overcome the wayfinding difficulties that people encounter in their everyday life due to the complexity of the urban environment. For many years, research related to location-based services has focused on outdoor navigation assistance. Aids that mainly answer to questions such as "what is the fastest way from my current location to the local hospital?" or "what is the optimal route from the train station to the University campus?". However, after arriving at a destination, using outdoor navigation entering a building to find the desired room is required. Indoor navigation assistance meets this need, by providing efficient indoor navigation instructions. The necessity for efficient and effective indoor navigation aids becomes even greater if we consider that according to research studies [1], people tend to lose orientation much easier within complex buildings, such as campus universities and hospitals.

The main problem of indoor navigation assistance approaches is the lack of accurate GPS signal, since it can only be used efficiently outside of buildings (radio signals cannot penetrate through walls). Due to this fundamental technical weakness, different approaches to overcome this problem were presented. These approaches can be based on the use of Wi-Fi, Bluetooth, ultra-wideband (UWB), near-field communication (NFC), etc.

Traditional static maps have been employed for many years as navigational assistance [2]. Nonetheless, indoor navigation aids nowadays, mainly rely on location-based services (LBS) [3–5]. In these approaches, aside from the major challenge of accurate positioning the user, communicating effectively the routing instructions is of great significance as well [6]. So far, several approaches have been proposed, informing the user about the optimal routing decision, facilitating in this manner the wayfinding procedure. Among these techniques, non-map-based approaches are gaining interest, mainly because they offer to the user the possibility of not missing important elements of the surrounding environment. Although indoor navigational assistance based on non-map-based approaches has gained popularity, studies have shown that the typical map-based approach is still the most efficient navigational aid for indoor environments [7,8]. Landmarks can play a significant role in human orientation and wayfinding. In fact, there are examples in literature [9] supporting that the enrichment of navigational assistance aids with landmarks can essentially improve navigation performance and user experience. In Duckham et al. [10] a category-based method for generating routing instructions based on landmarks was introduced. Although this approach aims on outdoor navigation, Fellner and her colleagues based their work on it and expanded it by developing an algorithm, the indoor landmark navigation Model, for an indoor navigation scenario [11].

This article aims to investigate the feasibility of the ILNM [11], by conducting a wayfinding experiment with human participants, as well as to combine it with indoor route maps and evaluate three different cartographic visualization scenarios.

The rest of the article is organized as follows. Section 2 presents related work. In Section 3, the proposed approach for investigating the feasibility of the ILNM is described. Section 4 presents the experiment procedure, while in the next section, we discuss the results obtained from this whole process. Section 6 discusses the results. In the last section, we draw conclusions and present possible future steps for expanding and improving the current work.

2. Related Work

2.1. Landmarks in Routing Instructions

Landmarks can be simply characterized as entities that are fixed in space and are at the same time useful for navigation purposes [12]. These entities are cognitively salient, quite prominent in the environment, and thus crucial in developing human spatial abilities (human spatial cognition). In Siegel and White [13], it is stated that understanding landmarks while someone interacts with a new environment is the first level of developing spatial knowledge. Although landmarks have been widely used in human wayfinding, there is an obvious lack of reference to them, regarding the procedure of generating routing instructions for a map service. The obvious explanation for this omission is the difficulty of developing a general pattern for identifying and then integrating landmarks within the routing instructions.

On the one hand, in the case of outdoor navigation, the identification and use of landmarks can be easily based on the prominent geometric, visual (e.g., colors), as well as semantic characteristics of the spatial objects that may serve as landmarks. Furthermore, in this case there is the advantage of having plenty of candidate spatial objects (e.g., buildings) that are suitable for landmarks. This type of information is also accessible in the form of geo-referenced images as well as in the databases of digital cadastral maps. On the other hand, there are a lot of difficulties when implementing the same approach in the case of indoor navigation. The main one has to do with the major structural differences

between indoor and outdoor spaces. Additionally, information regarding the indoor spatial features that might serve as landmarks is usually neither categorized nor available in most cases.

A lot of approaches have been proposed recently for generating landmark-based instructions. Raubal and Winter [9] first attempted to integrate landmarks in routing instructions for outdoor navigation. The whole approach was based on evaluating candidate spatial objects that may serve as landmarks, in terms of their visual, semantic, and structural saliency. In the later years, many researchers based their own scientific research on this approach. A quite different approach was presented by Duckham, in 2010 [10]. Duckham did not base his research on the semantic and structural characteristics of the candidate landmarks, but on the commonly available data about categories of landmarks, such as point of interests (POIs). This particular model (model of outdoor environments (OLNM)) consists of two main components:

i. Process of scoring the suitability of POIs that may serve as landmarks;
ii. Development of an algorithm for selecting and integrating the suitable landmarks within the routing instruction. This specific approach seemed suitable also in order to be applied on indoor environment

Summarizing, research regarding the use of landmarks within the routing instructions has shown that landmarks can be beneficial, in order the user to have a more pleasant user experience, as well as a more efficient and effective navigation performance. In a broader sense, it is evident that landmarks can improve the quality of routing instructions in terms of how successfully the user can interpret and finally use them.

2.2. Indoor Landmark Navigation Model (ILNM)

The indoor landmark navigation model (ILNM) is an algorithm for generating landmark-based route instructions to facilitate wayfinding activities in indoor environments [11]. Similar to the method proposed by Duckham et al. [10], this specific approach is not based on detailed instance-level data about the visual, semantic, and structural characteristics of individual spatial objects that may be used as landmarks. This method relies on using commonly available data about categories of spatial features, that can be found in most indoor spatial databases. The algorithm differentiates between indoor spatial objects and indoor spatial features. Indoor spatial objects represent real-world objects in an indoor environment, while indoor spatial features are abstractions or representations of indoor objects. An important aspect of the algorithm is that it makes use only of spatial features that can be represented in spatial databases. This relies on the fact that some of the indoor spatial objects in an indoor environment cannot be represented as spatial features in an indoor spatial database (e.g., a table).

The algorithm consists of three steps (Figure 1):

i.   Identification of categories of indoor spatial features that may serve as landmarks;
ii. Landmark selection for a specific path route from a set of candidate landmarks;
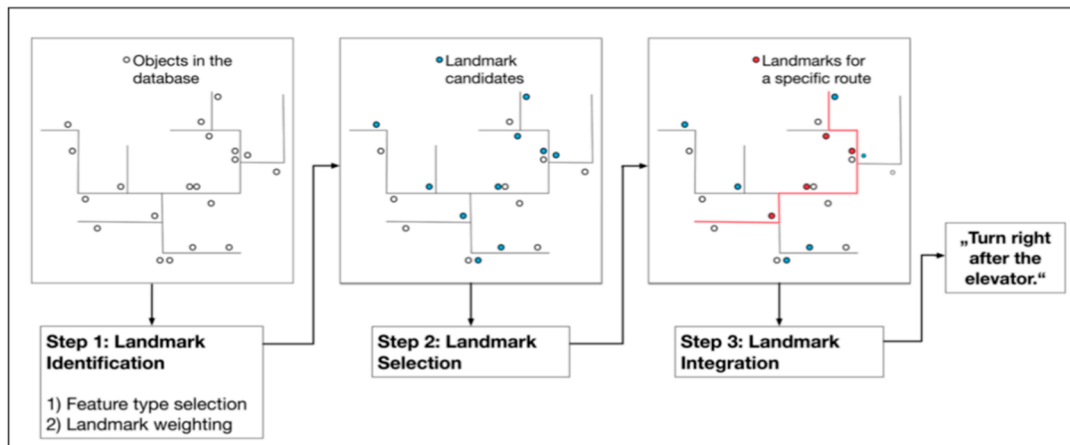iii. Landmark integration within the route instruction and generation of landmark-based instructions

Figure 1. Overview of indoor landmark navigation model; Source: [11].

3. Feasibility Investigation of ILNM (Indoor Landmark Navigation Model)

3.1. Selection of Test Area

The wayfinding experiment took place at ETH Zürich Hönggerberg campus and, more specifically, within the HIL Building. The starting point of the experiment was room E 19.1 (Figure 2) on the first floor, while the destination point was on the floor below, room D 55.2 (Figure 3). This specific route was selected for a variety of reasons:

i.   Along this route lie many spatial objects that may serve as landmarks. This element was quite crucial in selecting this specific path, since it was quite representative and suitable for applying the proposed algorithm [11].

ii. Furthermore, the fact of having along the route the element of changing from one level to the one below (from E to D floor), was important in the experiment's context, to fully investigate the feasibility of the algorithm in a real case scenario.

iii. The route consists of differences in length and in difficulty sub-paths. This element was important for selecting the specific routing path, since it represents a quite realistic scenario. In Figure 4, three different sub-paths of the wayfinding's routine path are presented. The paths differ in level of difficulty in terms of possible navigation decisions. For instance, in the first two paths, the users have to choose the right path, among other options. On the other hand, in the third path, the options are limited, making users' navigation decision much easier. In each of the presented cases (Figure 4), the existence of multiple and different types of landmarks can facilitate the whole navigation procedure.

3.2. Materials and Methods

3.2.1. Wizard of Oz Methodology

The navigation approach tested in this wayfinding experiment would eventually suffer from localization inaccuracies (e.g., lack of accurate GPS signal). Since the main aim of the wayfinding experiment was to investigate the feasibility of the ILNM algorithm resulting from the efficiency of the custom-made routing application, the "Wizard of Oz" methodology was applied [14]. This approach is mainly employed in cases where the need to avoid these type of technical weaknesses (e.g., signal inaccuracies) arises. It is a method used in experimental conditions, giving the participants the impression that they interact with the system, while in the same time the experimenter acts as "Wizard" and controls the interaction between the system and the participant.

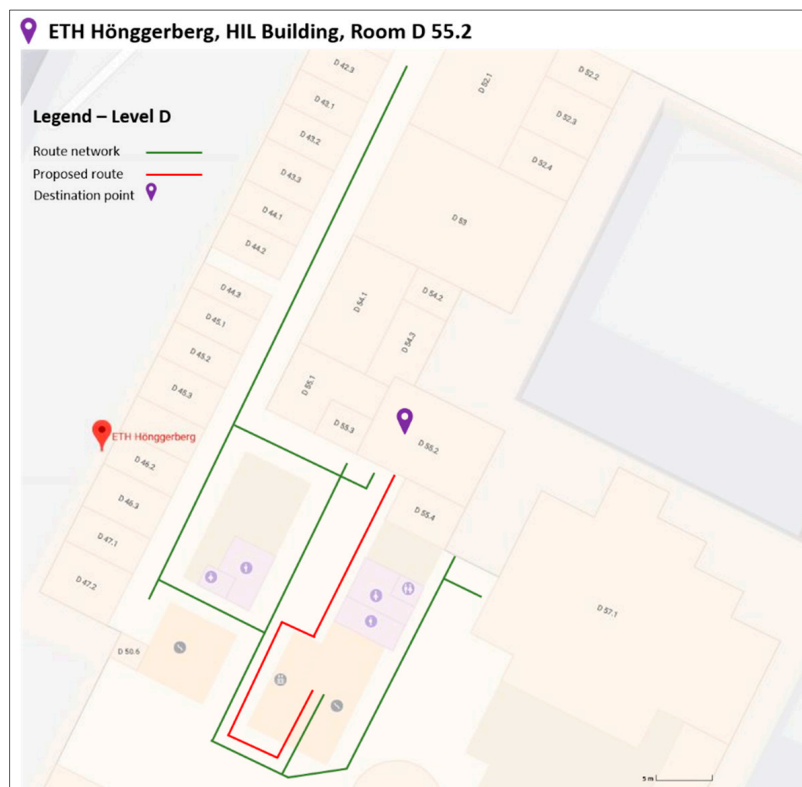**Figure 2.** Starting point (room E 19.1) and routing path on the E floor.

Figure 3. Destination point (room D 55.2) and routing path on D floor.

Figure 4. Three different sub-paths of the wayfinding's routing path.

In the experiment's framework, two android devices were used. One of them was used by the participant, just as a digital map, in which all the routing instructions were visualized. The second device was carried by the experimenter, in order to control the participant's device. In order to accomplish the aforementioned communication between the devices, many commercial applications were tested. The initial idea was the communication of the two devices through Wi-Fi. This approach was not ideal, mainly since the Wi-Fi signal within the experiment's area was not strong enough. Since the stability of the connection between the devices was of great significance for the success of the experiment, the communication based on Bluetooth signal between the devices was selected. The devices' connection became possible with the use of "Tablet Remote APK" application (Figure 5).
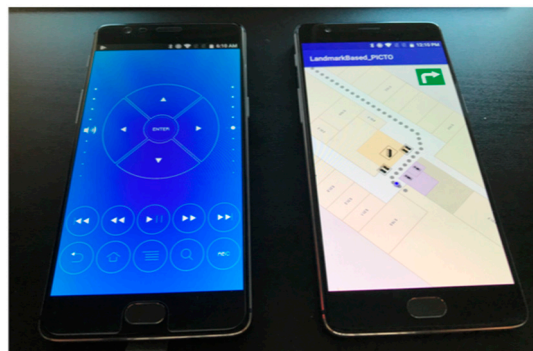

Figure 5. Two android devices connected via Tablet Remote APK (via Bluetooth).

3.2.2. Hardware Used

In the experiment's context, two identical android devices were used (Figure 6). In one of the devices, all the developed applications were loaded. This device was used during the experiment by the participants, as a digital map where all the routing instructions were visualized. The second device was used by the experimenter to remotely control the participant's device using the "Tablet Remote APK" Android application.

Figure 6. Two One Plus 3 devices, used in the wayfinding experiment.

In order to load the self-developed applications for the pre-testing process and the actual experiment to the mobile device, we had to enable the "developer options" in the smartphone's settings. Another adjustment that had to be made on the phone carried by the participants, was regarding the total time that the phone's screen would be "awake". Since during the experiment, any interaction with the phone was prohibited for the participants, it was essential to assure that the phone would remain "active" throughout the experiment's process. Based on this assumption, the display's "sleep" setting was adjusted to 30 min.

3.2.3. Software Used

In the context of designing the wayfinding experiment, three different prototype android applications were developed. Application development was based on Java programming language and the Android Studio integrated development environment (IDE). The basic components of these applications and the relevant software tools used to develop them, are the following:

- Base map (based on Google maps): The base map of the test area was designed in Adobe Photoshop. Based on this map, 85 to 87 (depending on each application) high-resolution static images of identical size were created. Each image had 2.5 m difference in the routing path from the next one (Figure 7). This change was vividly visualized by a blue circle, which was used as an indicator of participant's current location. The aforementioned design procedure was followed, in order to give the participants the impression of having a fully functioning routing application where the base map is dynamically updating based on the current location. Furthermore, within every image, the generated routing instructions were included. At the same time, indicators regarding the routing path (grey dots in Figure 7) as well as the proposed direction the participant had to follow, were designed and included in the final image (Figure 7).
- Java code: The application itself was written in the user interface of the Android Studio software (Figure 8).
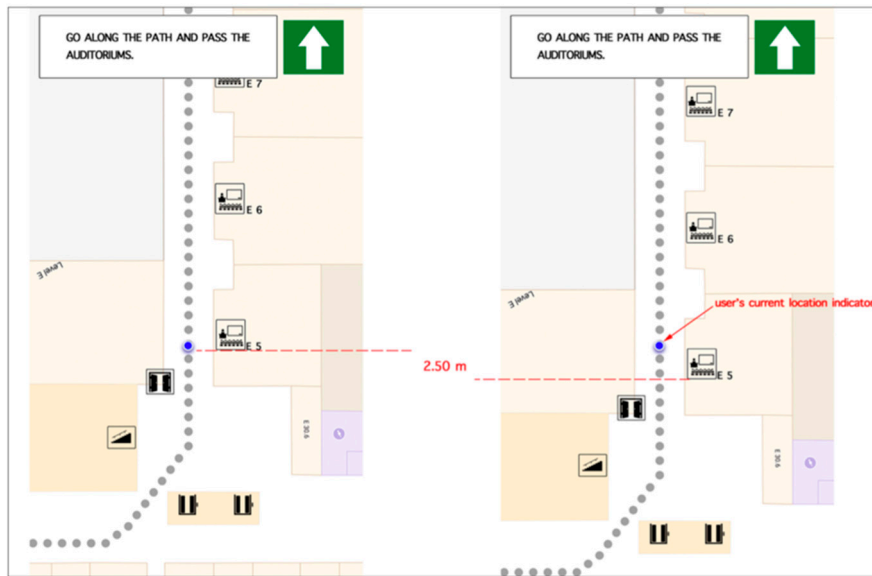
Figure 7. Successive images along the routing path (from left to right).

Figure 8. Java code of the prototype application.

The central part of the application was an image switcher, where all the generated images were loaded. The principal characteristic of an image switcher is that allows the addition of transitions on the images through the way they appear on screen. In the experiment's context, this element was quite crucial, since the way of switching from one image to another was essential for fulfilling the application's main aim.

An important aspect of the application's development was the design approach that was followed for the creation of the two buttons used by the experimenter, to switch from one image to another, while the participant was interacting with the application. On one hand, during the experiment, the user was not supposed to have any interaction with the phone and with the application in general. On the other hand, it was also essential to hide from the participant that the application was remotely controlled by the experimenter. Thus, the design of the buttons used in the routing application had to fulfil aforementioned assumptions. For that purpose, two large "invisible" buttons were designed.

The last step in the application's development was the creation of a method for saving to a .txt file inside android device's internal storage of the number of times the experimenter clicked the "Back" button during the experiment's procedure. This was important in the context of measuring every participant's navigation performance. The corresponding syntax for creative the aforementioned method is presented below (Figure 9).

```java
@Override
protected void onCreate(Bundle savedInstanceState) {(...}

// create function for saving info to external storage

private void saveText(String content) {

    //create file
    File file = new File(Environment.getExternalStorageDirectory().getAbsoluteFile(),
            FILENAME);

    //Write to file

    try {
        FileOutputStream fos = new FileOutputStream(file);
        fos.write(content.getBytes());
        fos.close();
    } catch (FileNotFoundException e) {
        e.printStackTrace();
    } catch (IOException e) {
        e.printStackTrace();
    }

}
```

Figure 9. Method for saving to a txt file within device's internal storage of the number of times "Back" button was clicked.

- Landmarks: One of the main aspects of the experiment's design was the visualization of the selected landmarks in two different forms. For this purpose, two design approaches were selected (pictogram-based and axonometric-based).

  ○ Landmarks in form of pictograms (Figure 10): In this approach, the design ideas were initially taken from www.flaticon.com. Based on these ideas, the final form of every landmark was further customized in Adobe Illustrator, which is a vector graphics editor developed and marketed by Adobe Systems.
  ○ Axonometric-based landmarks (Figure 11): Based on this design approach, every landmark was designed in ARCHICAD 21 (architectural BIM CAD).
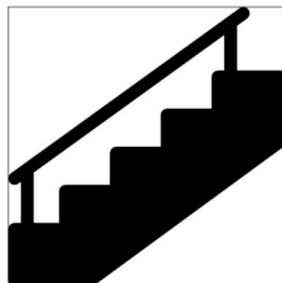

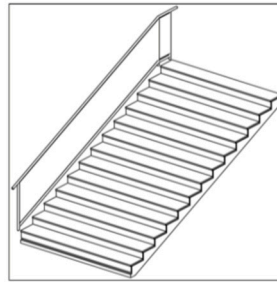**Figure 10.** Pictogram-based "Stairs".

Figure 11. Axonometric-based "Stairs".

3.3. Applying Indoor Landmark Navigation Model (ILNM)

The first step of the ILNM algorithm implementation was the identification of the indoor spatial features that could serve as landmarks. Only those feature categories whose instances were recognizable as well as available on the route were selected. In total, the following 15 spatial feature categories fulfilled the aforementioned condition.

Candidate landmarks		
1. Elevator	*7. Auditorium*	*13. Notice board*
2. Toilet	*8. Seminar room*	*14. Meeting room*
3. Door	*9. PC room*	*15. Trash and recycling can*
4. Scanner/Printer	*10. Lockers*	
5. Stairs	*11. Organic waste can*	
6. Closets	*12. Uncategorized room*	

The next step included the rating of landmarks suitability. The weighting procedure was carried out by a group of experts (3 females and 4 males), consisting of employees and students of ETH Zürich with at least 5 years of experience in "working" or "studying" in the HIL Building at Hönggerberg campus). During this procedure, the group of experts was asked to use a 5-point Likert scale and rate each of the candidate landmarks based on certain suitability factors with respect to the following two dimensions:

i.   How suitable a typical instance of this category is as a landmark (from "Ideal" to "Never suitable").

ii. How likely it is that a particular instance of this category is typical (from "All" to "Few").

For this purpose, a questionnaire was created in google forms and sent to the participants (Figure 12).

Figure 12. Questionnaire for landmarks suitability evaluation.

The scoring of the suitability for every spatial feature was based on the system proposed in Reference [10] (Figure 13).

Suitability	Typicality				
	All	**Most**	**Many**	**Some**	**Few**
Ideal	8	4	2	1	0
Highly suitable	4	4	2	1	0
Suitable	2	2	2	1	0
Somewhat suitable	1	1	1	1	0
Never suitable	0	0	0	0	0

Figure 13. Landmark scoring system based on spatial feature categories [10].

Below, in Table 1, an example rating for the spatial feature category of door is presented:

Table 1. Example ratings for the feature category "door".

Title	Suitability	Typicality
Physical size	Ideal	All
Prominence	Highly suitable	Most
Difference from surroundings	Highly suitable	Most
Availability of a unique label	Never suitable	All
Ubiquity and familiarity	Ideal	All
Length of description	Ideal	All
Spatial extents	Highly suitable	Most
Permanence	Ideal	All

The ratings above (Table 1) were then combined to generate an overall suitability score for the feature category "door", using the scoring system (Figure 13) proposed by Reference [10,11]. More specifically, the following score is obtained:

-   "Physical size" = 8, since it is rated as "Ideal" and as "All" in terms of suitability and typicality, respectively;
-   "Prominence" = 4, since it is rated as "Highly suitable" and as "Most" in terms of suit- ability and typicality, respectively;
-   "Difference from surroundings" = 4, since it is rated as "Highly suitable" and as "Most" in terms of suitability and typicality, respectively;
-   "Availability of a unique label" = 0, since it is rated as "Never suitable" and as "All" in terms of suitability and typicality, respectively;
-   "Ubiquity and familiarity" = 8, since it is rated as "Ideal" and as "All" in terms of suitability and typicality, respectively;
-   "Length of description" = 8, since it is rated as "Ideal" and as "All" in terms of suitability and typicality, respectively;
-   "Spatial extents" = 4, since it is rated as "Highly suitable" and as "Most" in terms of suitability and typicality, respectively;
-   "Permanence" = 8, since it is rated as "Ideal" and as "All" in terms of suitability and typicality, respectively.

Therefore, the final score for the category of door is computed as follows: 8 + 4 + 4 + 0 + 8 + 8 + 4 + 8 = 44.

To assign a score for the suitability of every candidate landmark, a python script (Figure 14) was created based on the proposed landmark scoring system [10].

```python
# in terms of PHYSICAL SIZE
suitSize = int(input('enter your suitability evaluation in terms of physical size: '))
typSize = int(input('enter your typicality evaluation in terms of physical size: '))

if suitSize == 5 and typSize==5:
    score1 = 8
if suitSize == 5 and typSize==4:
    score1 = 4
if suitSize == 5 and typSize==3:
    score1 = 2
if suitSize == 5 and typSize==2:
    score1 = 1
if suitSize == 5 and typSize==1:
    score1 = 0

# suitability highly suitable

if suitSize == 4 and typSize==5:
    score1 = 4
if suitSize == 4 and typSize==4:
    score1 = 4
if suitSize == 4 and typSize==3:
    score1 = 2
if suitSize == 4 and typSize==2:
    score1 = 1
if suitSize == 4 and typSize==1:
    score1 = 0

# suitability suitable
if suitSize == 3 and typSize==5:
    score1 = 2
if suitSize == 3 and typSize==4:
    score1 = 2
if suitSize == 3 and typSize==3:
    score1 = 2
if suitSize == 3 and typSize==2:
    score1 = 1
```

Figure 14. Python script for scoring the suitability of spatial features based on the algorithm proposed in Reference [10].

After all the candidate spatial feature categories were rated and scored, a normalized weight for every spatial feature was calculated, which resulted in the following list of landmarks (Table 2).

3.3.1. Landmark Selection

The selection of the landmarks that were integrated within the routing instructions was based on certain parameters presented in Reference [11]. In this context, the adjustment unit (*au*) that was used, for any decrease or increase of the suitability weight for every landmark, was equal to 0.2. As it is proposed Reference [11], this value represents a fifth of the maximum initial weight. It is a value that will certainly provide an effect on the final weight, without distorting significantly the initial weight resulted from the weighting process.

The landmarks that were finally selected are the following six (Table 3).

Table 2. Overall suitability scores and normalized weights of the selected landmarks.

Landmarks	Suitability Score	Normalized Weight
Elevator	35	1.00
Auditorium	30	0.81
Stairs	29	0.80
Toilet	26	0.67
Seminar Room	24	0.59
Lockers	23	0.56
Door	22	0.55
PC Room	18	0.38
Trash and Recycle Can	17	0.37
Meeting Room	15	0.28
Scanner/printer	15	0.27
Organic waste can	14	0.26
Meeting Room	15	0.28
Notice Board	13	0.22
Uncategorized room	7	0.00

Table 3. Final selected landmarks and their corresponding normalized weights.

Landmarks	Suitability Score	Initial Weight	Final Weight
Elevator	35	1.00	**1.00**
Toilet	26	0.67(+0.2)	**0.87**
Stairs	29	0.80	**0.80**
Door	22	0.55(+0.2)	**0.75**
Auditorium	30	0.81−[(5−1)*0.2/4]	**0.61**
Scanner/Printer	15	0.27(+0.2)	**0.47**

In the table above, the final weights of the selected landmarks are presented. The final weight of every of the six landmarks is obtained based on certain parameters that are included in the algorithm presented in Reference [11]. For instance, we increased the weight of the spatial features of "toilet", "door", and "Scanner/Printer" based on their position along the route. More specifically, these three landmarks were located before certain decision points that lie on the route. This specific element increased the significance of the landmarks, something that is illustrated by increasing their weight by the adjustment unit (*au*). Similarly, the final weight of the spatial feature "Auditorium" was decreased by 0.2. This unit was obtained based on the following formula: n−1*(au/4). According to Reference [1], when along the same route leg multiple instances of the same landmark category occur, then their final weight has to be set to zero when their count *n* exceeds five, otherwise their suitability weight has to be decreased by the formula presented above [n−1*(au/4)]. In our case, four instances of the landmark category "Auditorium" occur along a certain route leg.

3.3.2. Generating Landmark-Based Routing Instructions

The final step of the ILMN algorithm is the generation of the landmark-based routing instructions (with the integration of the selected landmarks). The basic principles for this procedure are presented in the Section 2.2. Based on the principles for generating landmark-based routing instructions presented in Reference [11], the outcome in our case was the following:

- "Turn right and pass through the door";
- "Go along the path";
- "Go along the path and pass women's toilet E 20.5";
- "Pass through the door and turn left";
- "Go along the path. You will pass through one door";
- "After the men's toilet E 11.2 turn right";
- "Go along the path. After the scanner/printer, turn slight left";
- "Go along the path and pass the stairs on your left";
- "Go along the path and pass the auditoriums";
- "After the auditorium E9, use the stairs to go to the level below";
- "Turn right";
- "Turn right and pass the elevators on your right";
- "Go along the path. You will pass through one door";
- "The path leads straight to your destination point".

In the experiment's context, for comparison reasons (metric-based application development), the corresponding metric-based instructions that were used in the benchmark approach for indoor navigation tasks were generated:

- "After 2.5 m turn right";
- "Go along the path and after 7 m turn left";
- "Go along the path. After 33 m turn right";
- "Go along the path. After 25 m turn slight left";
- "Go along the path for the next 67 m";
- "Use the stairs to go the level below";
- "Turn right";
- "Go along the path and after 7 m turn right";
- "Go along the path for the next 24 m";
- "The path leads straight to your destination "room D 55.2"".

3.3.3. Design Procedure of the Selected Landmarks

Pictogram-Based Landmarks

The ideas for visualizing the landmarks in this design form were taken by www.flaticon.com. These initial ideas were further customized within Adobe Illustrator. Below (Figure 15), the final pictogram-based version of the selected landmarks is presented.
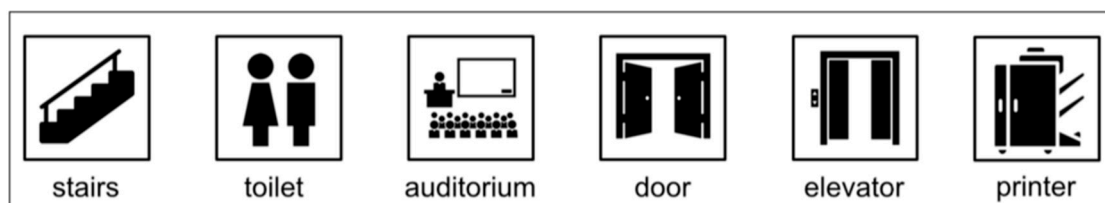

Figure 15. Pictogram-based landmarks.

Axonometric-Based Landmarks

The selected landmarks, based on the axonometric design approach, were created within ARCHICAD 21 (architectural BIM CAD). This final version of the landmarks is presented below (Figure 16).
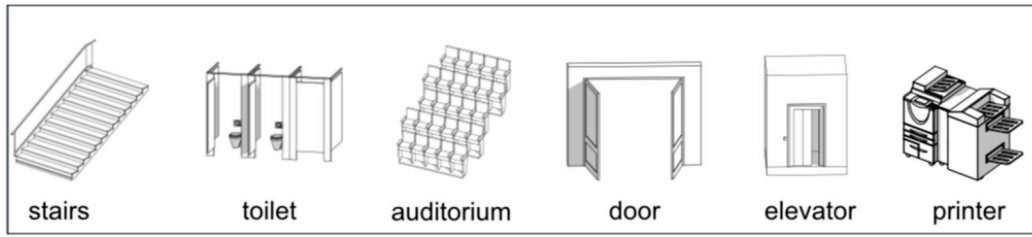
Figure 16. Axonometric-based landmarks.

4. User Study

To investigate the feasibility of the ILNM, a user study was designed. The study consisted of the following three discrete components:

i.   A pilot study;
ii.   Wayfinding experiment with 30 participants;
iii.   Results interpretation.

4.1. Pilot Study

In order to resolve basic issues of the system already before the actual experiment, a pilot study was conducted, with a usability expert testing the system and providing feedback. The main points of his feedback were the following:

- A pre-test application would be beneficial to improve the quality of the collected data during the final experiment. Participants would have the opportunity to get familiarized with the routing application, thus they could test and give feedback regarding the application in a more efficient and unbiased way.

The axonometric-based landmark visualization seemed confusing, since each landmark visualization is quite detailed. The proposal was to integrate each landmark within the framework, where the routing instructions are visualized (Figure 17).
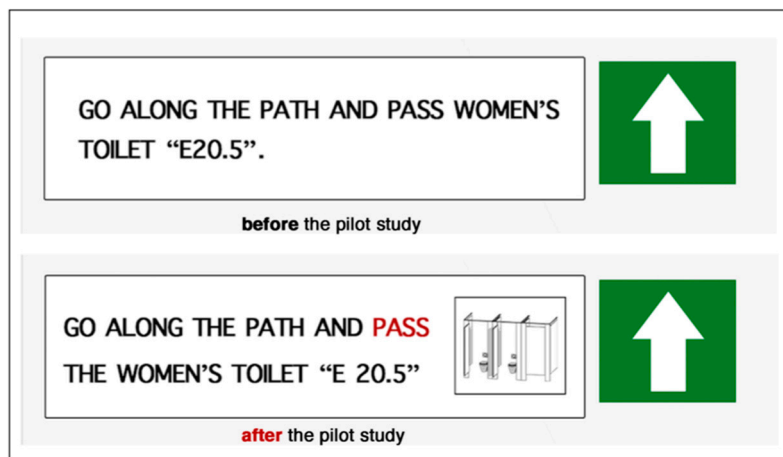

Figure 17. Integration of axonometric-based landmarks within the routing instructions (based on pilot study feedback).

4.2. Experiment Procedure

At the beginning of the experiment, the participants were informed about the conditions and the steps of the experiment. More specifically, the experiment included the following procedures:

   i.   Two pre-study questionnaires. In the first one, the participants were asked to provide their demographic information (age, country of origin, current profession), their level of experience with digital maps and navigation systems, as well as their experience in navigating within the HIL building (test area). The second questionnaire was the Santa Barbara sense of direction scale (SBSODS), a standardized questionnaire for the self-report of participants' spatial abilities [15].

   ii.   Participation in a small pre-test experiment in order the participants to get familiarized with the indoor navigation assistance application.

   iii.   Participation in the actual experiment (Figure 18).

   iv.   Answer to three questionnaires regarding participants' overall user experience, the cognitive workload of theirs, as well as the evaluation of the design aspects of the application (Figure 19).

Figure 18. Participant and experimenter during the wayfinding experiment.
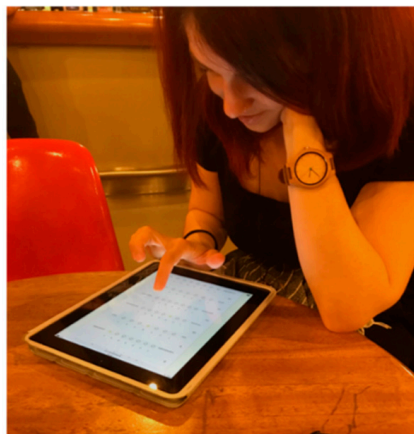
Figure 19. Participant fills in the questionnaire, after finishing experiment's wayfinding task.

4.3. Tracked Data

During the experiment, the participants were asked to complete the navigation task, using only the routing instructions that were visualized on the phone's screen. Any interaction with the phone device (e.g., click, zoom in-out) was strictly prohibited. At the same time, the experimenter was walking alongside the participants, controlling their phones, based on the "Wizard of Oz" methodology. Furthermore, the total time that every participant needed to complete the experiment's task was recorded, with the use of "stopwatch app" in experimenter's phone. Another important parameter regarding the participant's navigation performance was the number of errors they made, during the navigation task. Every time the participants were choosing to navigate towards the wrong direction, meant that the experimenter needed to press the "Back button" in order to readjust the map. This

action (press of the back button) was "translated" within the java code as "the user made an error". This information was saved as a .txt file within the Androids phone's internal memory.

4.4. Participants

In total, 30 participants volunteered for the experiment, 10 for each of the three conditions. The 10 participants of the 'pictogram' condition (8 males) had a mean age of 28.9 years (SD = 4.43). Similarly, for the 'axonometric' condition, the 10 participants (9 males) had a mean age of 30.1 years (SD = 9.66). The 10 participants of the 'metric' condition (6 males) had a mean age of 28.4 years (SD = 5.36). Furthermore, in order to ensure that there were no significant differences among the participants' characteristics in the three scenarios, statistical analyses of the participants' answers concerning their demographic data, level of experience with navigation systems and digital maps, as well as their spatial abilities, were performed. For this purpose, a Kruskal–Wallis *H* Test was run on the data collected from the two pre-study questionnaires. The Kruskal–Wallis *H* Test showed that there was no statistically significant difference in any of the participants' answers (Table 4).

Table 4. Kruskal–Wallis *H* Test results regarding pre-study participants' answers, for the three conditions.

	Age	Experience with Digital Maps	Experience with Navigation Systems	How Many Hours did you Sleep Last Night?	How do You Feel?
Kruskal-Wallis H	0.334	0.460	0.745	0.853	4.896
df	2	2	2	2	2
Assymp. Sig.	0.846	0.794	0.689	0.653	0.086

The same procedure was followed for the participants' answers regarding their spatial abilities. According to the table below (Table 5), there was no statistically significant difference in participants' spatial abilities among the three conditions ('pictogram', 'axonometric', 'metric').

Table 5. Kruskal–Wallis *H* Test results regarding pre-study participants' spatial abilities.

	Santa Barbara Sense of Direction Scale
Kruskal-Wallis H	2.313
Df	2
Assymp. Sig.	0.315

4.5. Hypothesis Statement

Based on the experiment's design aspects and conditions, the following hypothesis were stated:

- H1: The navigation performance based on the two landmark-based routing applications will be more efficient in terms of number of errors occurred during the navigation task, compared to the benchmark approach (metric-based).
- H2: The navigation performance based on the two landmark-based routing applications will be more efficient in terms of total completion time of experiment's navigation task, compared to the benchmark approach (metric-based).
- H3: The user experience in terms of attractiveness, of the two landmark-based routing applications (pictogram-based and axonometric-based), will be better, in comparison to the benchmark approach (metric-based).
- H4: The user experience in terms of pragmatic quality aspects (perspicuity, efficiency, and dependability), of the two landmark-based routing applications (pictogram-based and axonometric- based), will be better, in comparison to the benchmark approach (metric-based).

- H5: The overall user experience in terms of hedonic quality aspects, of the two landmark-based routing applications (pictogram-based and axonometric-based), will be better, in comparison to the benchmark approach (metric-based).
- H6: The perceived cognitive workload in terms of effort and frustration, will be lower during the landmark-based approaches, in comparison to the benchmark approach (metric-based).
- H7: The perceived cognitive workload in terms of performance, will be higher during the landmark-based approaches, in comparison to the benchmark approach (metric-based).
- H8: The overall perceived cognitive workload will be higher during the benchmark approach (metric-based).

The two first hypotheses (H1 and H2) are based on the fact that the two landmark-based routing applications will convey the routing instructions to the users in a much more unambiguous and clear way, helping them to have an overall more effective and efficient navigation performance. The next three hypotheses (H3, H4, and H5) are expected to be confirmed, due to the clearer and much more content-enriched user interface of the landmark-based approaches, in comparison to the benchmark approach (metric-based). Hypothesis (H6) is based on the fact that the existence of landmarks, not only within the routing instructions, but also on the base-map as visualizations, will be beneficial for the users in order to position and navigate themselves easier within the test area. This fact will help them to complete the wayfinding task with the least possible effort and frustration. Hypothesis (H7) is based on the overall impression of security that the two landmark-based approaches are expected to provide to the users, in comparison to the benchmark approach (metric-based). The final hypothesis (H8) is expected to be confirmed, based on the previous hypotheses (H7 and H6).

5. Statistical Analysis and Graphs

The experiment's data analysis was conducted for the categories of navigation performance, user experience (UX), cognitive workload, as well as for best suitable form of landmarks visualization. In terms of statistical analyses, non-parametric tests were used, since the dependent variables are not normally distributed. More specifically, for comparison among the three different conditions (pictogram, axonometric, metric) a Kruskal–Wallis *H* Test was performed. Furthermore, for pairwise comparisons (axonometric vs. pictogram, axonometric vs. metric, pictogram vs. metric), the non-parametric Mann–Whitney tests were utilized.

5.1. Navigation Performance

One of the main aspects of this experiment was to investigate how efficiently the participants completed the navigation task, in all three cartographic scenarios. The navigation performance in the three conditions was analyzed by measuring the completion time in each case as well as the number of errors that occurred during the navigation task. The descriptive statistics (SD = standard deviation and mean value) for all scenarios are depicted in the table below (Table 6).

Table 6. Descriptive statistics for number of errors and completion time.

	Number of Errors_Mean	Completion Time (min)_Mean
Pictogram	0.40	2.94
Axonometric	0.40	2.85
Metric	2.30	2.80
	Number of Errors_SD	Completion Time (min)_SD
Pictogram	0.69	0.29
Axonometric	0.84	0.48
Metric	1.34	0.41

The next step was to investigate whether there were statistically significant differences, in terms of errors during the navigation task, as well as in terms of the total completion time of navigation task, among the three conditions (1: Pictogram, 2: Axonometric, 3: Metric). For this purpose, a Kruskal–Wallis *H* Test was utilized. The specific test was selected because it is a rank-based non-parametric test that can be used to determine if there are statistically significant differences between two or more groups of an independent variable on a continuous or ordinal dependent variable. In our case, all the aforementioned conditions were met, since there are, in total, three independent variables (1: Pictogram, 2: Axonometric, 3: Metric) related to two continuous variables (1: Number of errors, 2: Completion time). Before actually performing the test, one more assumption had to be met, in order to get valid results. This assumption is related to the independence of the variables, which means that there should not be any relationship between the observations in each group of the independent variable or between the groups themselves. In the experiment's context, this assumption was also met. In the table below (Table 7), the results from the Kruskal–Wallis H Test are presented.

Table 7. Kruskal–Wallis H Test results of navigation task's completion time and number of errors among three scenarios (1: Pictogram, 2: Axonometric, 3: Metric).

	Scenario	N	Mean Rank
Completion Time	Pictogram	10	17.55
	Axonometric	10	14.95
	Metric	10	14.00
	Total	30	
Number of errors	Pictogram	10	13.60
	Axonometric	10	12.90
	Metric	10	20.00
	Total	30	
	Completion Time	**Number of Errors**	
Kruskal-Wallis H	0.873	5.102	
df	2	2	
***p*-value**	0.646	0.78	

In a next step of the analysis, a Mann–Whitney *U* test was performed for the two continuous variables (completion time and number of errors), between all pairs of conditions (pictogram vs. axonometric, pictogram vs. metric, axonometric vs. metric). The Mann–Whitney *U* test (also called the Wilcoxon–Mann–Whitney test) is a rank-based nonparametric test that can be used to determine if there are differences between two groups on a continuous or ordinal dependent variable. This is the main reason why this test is utilized in pairwise comparison cases, instead of performing a Kruskal–Wallis *H* test again.

In our analysis context, the Mann–Whitney *U* test was utilized to extend the analysis of our data in a pairwise condition. Although the results from the Kruskal–Wallis H test (Table 7) did not reveal any statistically significant differences, this test was carried mainly to identify possible differences for the continuous variable of "number of errors" (*p*-value = 0.78). Below in Table 8, the corresponding results are presented.

Table 8. Mann–Whitney U test results between all pairs of conditions (pictogram vs. axonometric, pictogram vs. metric, axonometric vs. metric).

	Completion Time	Number of Errors
Mann–Whitney U (Pictogram vs. Axonometric)	41.50	47.00
Wilcoxon W	96.50	102.00
Z	−0.643	−0.299
p-value	0.520	0.765
Mann–Whitney U (Axonometric vs. Metric)	47.00	27.00
Wilcoxon W	102.00	82.00
Z	−0.227	−1.915
p-value	0.820	0.055
Mann–Whitney U (Pictogram vs. Metric)	34.00	20.50
Wilcoxon W	89.00	65.50
Z	−0.899	−2.203
p-value	0.369	**0.028**

5.2. User Experience

The assessment of the user experience of the three cartographic scenarios was based on the User Experience Questionnaire (UEQ) [16]. The UEQ is used to assign a score to each of the approaches for the following six user experience categories [17]:

- *Attractiveness*: Overall impression of the product. Do users like or dislike it? Is it attractive, enjoyable, or pleasing?
- *Perspicuity:* Is it easy to get familiar with the product? Is it easy to learn? Is the product easy to understand and unambiguous?
- *Efficiency*: Can users solve their tasks without unnecessary effort? Is the interaction efficient and fast? Does the product react to user input quickly?
- *Dependability*: Does the user feel in control of the interaction? Can he or she predict the system's behavior? Does the user feel confident when working with the product?
- *Stimulation*: Is it exciting and motivating to use the product? Is it enjoyable to use?
- *Novelty*: Is the product innovative and creative? Does it capture the user's attention?

According to Reference [17], the aforementioned user experience categories are not considered to be independent. For instance, user's overall impression is recorded by the attractiveness scale, which at the same time is influenced by the other five scales. More specifically, the six user experience scales can be further categorized as follows:

- *Attractiveness*: Pure dimension that indicates emotional reaction on a pure acceptance/ rejection context.
- *Pragmatic quality aspect*: User experience scales describing interaction qualities, related to tasks or goals the user aims to reach while using the product. In this category belong the scales of perspicuity, efficiency, and dependability.
- *Hedonic quality aspect*: Scales that are not related to tasks and goals, but with user experience aspects, associated with pleasure and satisfaction. In this category belong the scales of stimulation and novelty.

The results based on the participant's answers, are presented in Figures 20 and 21. Furthermore, for a clearer and overall better presentation of the user experience scores, a comparison with a benchmark dataset is presented in Figure 22 [18]. In this figure, the differences also among the tested scenarios (pictogram, axonometric, metric) are clearly illustrated.
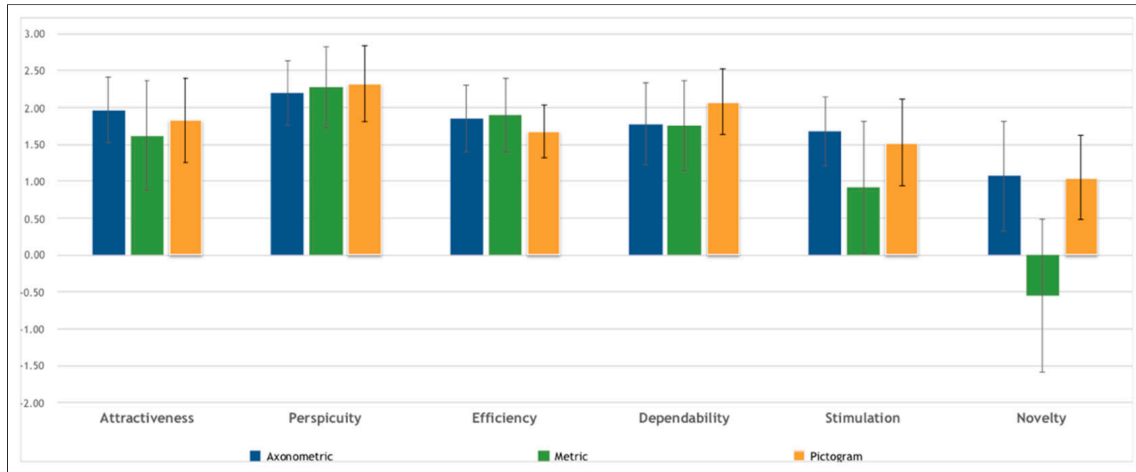

Figure 20. User experience results for each user experience (UX) category and condition. The black error bars indicate the confidence level.
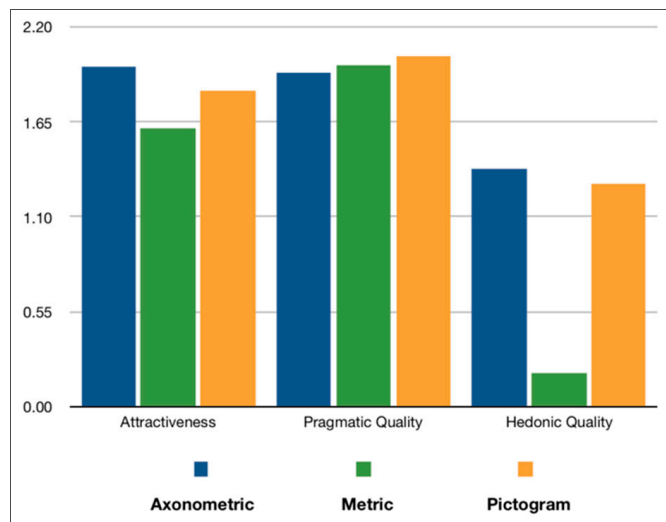

Figure 21. User experience results for every condition, in terms of attractiveness, pragmatic, and hedonic quality aspects.

In a further statistical analysis, a Kruskal–Wallis *H* test was utilized, in order to examine the existence of statistically significant differences in the six user experience categories, among the three different conditions (pictogram, axonometric, metric). In the table below (Table 9) the results are presented.

Table 9. Kruskal–Wallis H test results for each of the six user experience categories, for all three conditions (pictogram, axonometric, metric).

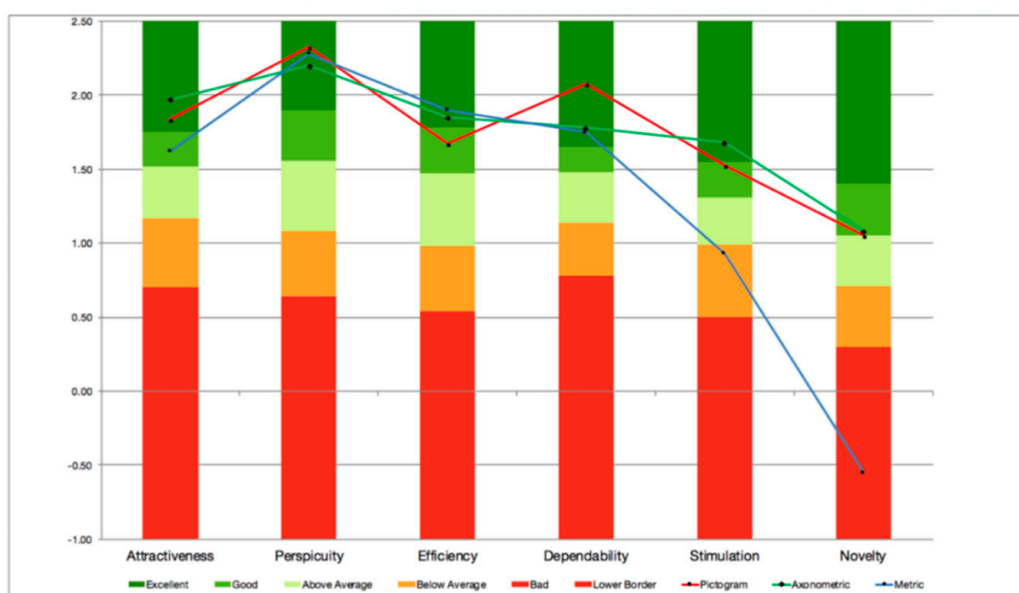|	Attractiveness	Perspicuity	Efficiency	Dependability	Stimulation	Novelty
Kruskal–Wallis H	0.332	0.411	0.555	0.570	1.614	6.846
df	2	2	2	2	2	2
***p*-value**	0.851	0.814	0.758	0.752	0.446	**0.033**

Figure 22. User experience results per category (attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty) and per condition, in comparison with a benchmark dataset.

5.3. Cognitive Workload

To measure and compare the cognitive workload that was required while using the different approaches, the raw NASA Task Load Index (TLX) questionnaire was utilized [19]. Based on the participants' answers, cognitive workload estimates for six categories were measured:

1: Mental demand: "It measures how much mental and perceptual activity was required, as well as how demanding, simple, or complex the task was.";

2: Physical demand: "It measures how much physical activity was required, as well as how easy or demanding, slack or strenuous the task was.";

3: Temporal demand: "It measures how much time pressure the participant felt due to the pace at which the tasks or task elements occurred.";

4: Performance: "It measures the level of success the participants felt in terms of completing the task, as well as how satisfied they were with their performance.";

5: Effort: "It measures the level of work (mentally and physically) the participants had to put in accomplishing their level of performance.";

6: Frustration: "It measures how irritated, stressed, and annoyed versus content, relaxed, and complacent the participants felt during the task.".

All participants were asked to answer six questions using a scale with 21 gradations starting from 1 that meant 'very low' to 21 that meant 'very high'. The gradations had a different meaning in the 'performance' question, since 1 meant 'perfect' and 21 meant 'failure'. Below, in Figure 23, the results for every category and scenario are presented.

Based on the figure above (Figure 23), there was no clear difference among the three approaches (axonometric, metric, and pictogram) in any of the six cognitive workload categories. For that reason, in order to analyze in depth and in pairwise conditions possible statistically significant differences, a Mann–Whitney *U* test was utilized. Below in Table 10, all the results are presented.
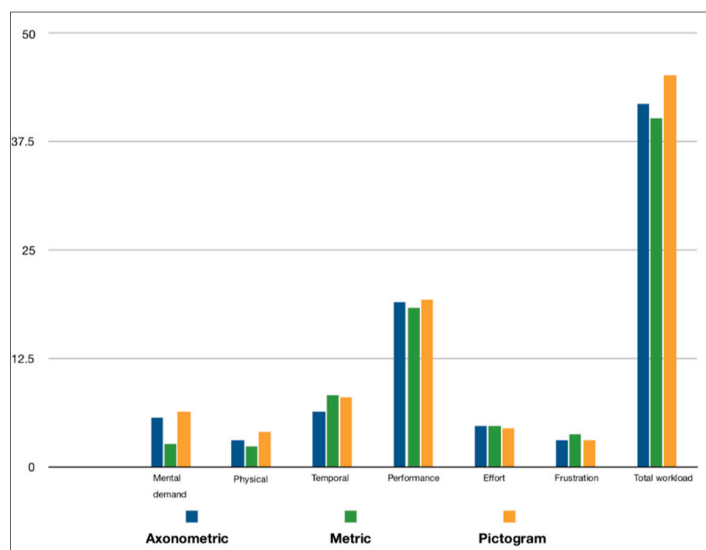
Figure 23. Cognitive workload estimates for six categories, for all the three conditions. (1. axonometric; 2. metric; 3. pictogram).

Table 10. Mann–Whitney U test results for all cognitive workload categories, between all pairs of conditions (pictogram vs. axonometric, pictogram vs. metric, axonometric vs. metric).

	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration	Total Workload
Mann–Whitney *U* (Pictogram vs. Axonometric)	48.00	38.50	36.00	49.50	46.00	41.50	43.50
Wilcoxon W	103.00	93.50	91.00	104.50	101.00	96.50	98.50
p-value	0.875	0.344	0.269	0.965	0.753	0.461	0.621
Mann–Whitney *U* (Pictogram vs. Metric)	28.50	31.00	46.00	41.50	47.50	40.50	41.00
Wilcoxon W	83.50	86.00	101.00	96.50	102.50	95.50	96.00
p-value	0.083	0.112	0.749	0.476	0.844	0.413	0.490
Mann–Whitney *U* (Axonometric vs. Metric)	25.00	43.00	39.50	43.00	49.50	47.00	47.00
Wilcoxon W	80.00	98.00	94.50	98.00	104.50	102.00	102.00
p-value	**0.045**	0.547	0.401	0.558	0.967	0.804	0.815

5.4. Landmarks Visualization Design Evaluation

As a final step in the experiment's context, all 30 participants were asked to evaluate which would have been the best design approach of the selected landmarks' visualization. More specifically, they had to choose among three different design approaches:

- Pictogram-based;
- Axonometric-based;
- Text-based.

The purpose of this evaluation process was for the participants to compare and assess different types of landmarks cartographic visualizations. A quite interesting aspect of this procedure was the fact that they had to evaluate, not only the approach they tested during the task, but also the two other possible visualization approaches. The evaluation procedure was included in the last questionnaire that the participants had to fill in as a last step of the experiment (Figure 24).
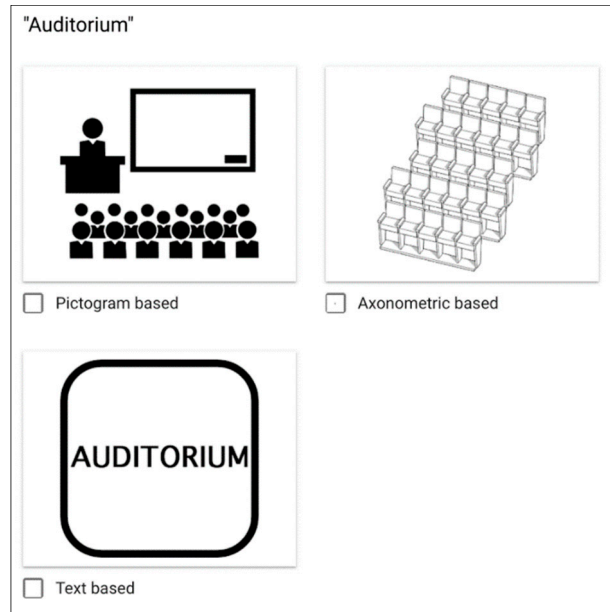
Figure 24. Evaluating different design approaches (1. pictogram-based; 2. axonometric-based; 3. text-based) for the landmark "Auditorium".

In the figure below (Figure 25), the results regarding participants' preferences on the best design approach for each of the five landmarks, are presented in the form of pie charts.
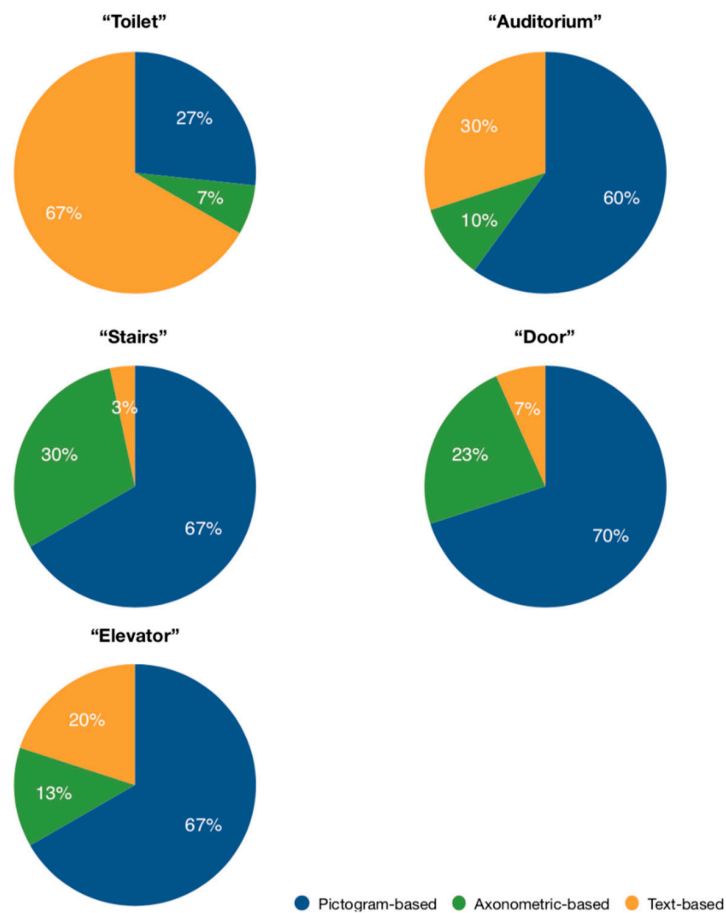

Figure 25. Participants' preferences regarding landmarks' visualization design approach.

6. Discussion

6.1. Navigation Performance

In the experiment's framework, participants' navigation performance was measured in terms of two dimensions:

i.   Total number of errors occurred during the navigation task;
ii.   Total time each of the participants needed to complete the navigation task.

Based on the results presented in the previous chapter (Table 6), in terms of efficiency, the two landmark-based routing applications are in line with our initial hypothesis (H1). The mean number of errors that occurred while participants were testing the two landmark-based approaches (mean number of errors = 0.40) was much smaller than the one that occurred during the testing of the benchmark approach (mean number of errors = 2.30). On the other hand, the total time needed for every participant to complete the navigation task was less in the case of the benchmark approach, in comparison to the two landmark-based approaches. This result contradicts what is stated in the second hypothesis (H2).

For the statistical analysis, a Mann–Whitney U test was performed between all pairs of conditions. The test revealed significant differences in terms of numbers of errors for the pairwise comparison of the pictogram and metric approaches (Table 8). More specifically, the number or errors occurred during the testing of the metric approach was statistically significantly higher than the number of errors during the pictogram approach, $U = 20.50$, $z = -2.203$, $p = 0.028$.

The results above indicate the positive effect of landmark-based routing instructions on the efficiency of participants' navigation performance. Even though the total time of completion of experiment's task was quite similar in all three conditions, the fact that significantly less errors occurred during the landmark-based approaches leads to the conclusion that the integration of landmark-based instructions within a routing application using ILNM can be very beneficial for the overall navigation performance.

6.2. User Experience

The analysis of the results gathered from the UEQ questionnaire confirmed the third hypothesis (H3), since the overall score of the two landmark-based approaches in terms of the user experience category of attractiveness was higher in comparison to the metric-based approach (Figures 20 and 21). The specific result is of great significance, since the specific UX category (attractiveness) is a strong indication of rejection or approval of an approach.

The fourth hypothesis (H4) stated the expectation that the user experience in terms of pragmatic quality aspects (perspicuity, efficiency, and dependability) of the two landmark-based approaches (pictogram and axonometric) would be better in comparison to the benchmark approach (metric-based). However, the results showed that the best approach in terms of UX pragmatic quality aspects was the pictogram-based approach. The second best was the benchmark approach (metric-based), while among the three, the worst one was the axonometric-based approach (Figure 21). Thus, this hypothesis (H4) was rejected. One possible explanation for this result might be the fact that, in the UX category of perspicuity, the axonometric-based approach received the lowest score among the three approaches. The existence of axonometric-based landmarks, which can be characterized as too detailed or even confusing, could led to this low score.

The fifth hypothesis (H5) was related to the UX hedonic aspects (stimulation and novelty) of the three approaches. The specific hypothesis was confirmed, since the overall score of the two landmark-based approaches (pictogram and axonometric) was significantly higher than the one the benchmark approach (metric-based) received. Hedonic quality aspects are UX categories that are not related to tasks and goals, but with user's pleasure and satisfaction. Novelty is one of UX scales that belong to this category. As it is vividly illustrated in Figure 21, the metric-based approach received

significantly lower ratings in this specific scale. This fact is further confirmed through the statistical analysis (Kruskal–Wallis H test), that was performed for each of the six user experience categories (attractiveness, perspicuity, efficiency, dependability, stimulation, novelty), for all three conditions (pictogram, axonometric, metric). More specifically, statistically significant differences for the UX category of 'novelty', among the three approaches, were observed (χ2(2) = 6.846, *p* = 0.033) (Table 9).

6.3. Cognitive Workload

In terms of measuring the perceived cognitive workload, the raw NASA TLX questionnaire was utilized [19]. The first hypothesis related to the perceived cognitive workload (H6) is partially in line with the results obtained from participants' answers. On the one hand, the participants did not need to put less effort into navigating, while they were testing the two landmark-based approaches. In fact, the level of effort required by the participants, was almost the same among the three approaches. On the other hand, as it was stated and expected in (H6), the participants experienced higher level of frustration, while they were testing the metric-based approach, in comparison to what they experienced during the testing of the two landmark-based approaches (Figure 23).

The hypothesis (H7) stated the expectation that the perceived cognitive workload in terms of performance would be higher for the participants during the two landmark-based approaches. The results show (Figure 23), that the hypothesis (H7) was confirmed. Although the total difference among the three approaches was not significant, the results indicate that the landmark-based approaches compared to the metric-based approach, provided to the participants a higher level of certainty regarding the successful accomplishment of the navigation task. A possible explanation for that could be the fact that the existence of landmarks within the routing instruction provide a sense of confirmation, as far as the correctness of the selected navigation path is concerned.

According to the final hypothesis (H8), the total perceived cognitive workload was expected to be higher for the participants, during the testing of the metric-based approach. However, the results show that the total perceived cognitive workload for the participants during the metric-based approach was lower, compared to the workload perceived during the other two approaches (Figure 23). As it is illustrated in Figure 23, the cognitive workload perceived during the metric-based approach, in terms of mental demand, was clearly lower in comparison to the two landmark-based approaches. This fact is further confirmed, since according to the statistical analysis (Mann–Whitney *U* test), a statistically significant difference in the pairwise comparison between the metric and axonometric approaches was observed, *U* = 25.00, z = −2.002, **p = 0.045**. One possible explanation for this observation could be the difference in the way of conveying the routing instruction among the three approaches. Landmark-based instructions require user's attention to try to match the displayed information on the screen with real spatial information in the environment. This is something that can push the user harder, mentally as well as physically, to achieve an adequate performance level during the navigation task.

Summarizing, according to the statistical analysis performed on the participants' answers (Mann–Whitney *U* test for all cognitive workload categories, between all pairs of conditions) in terms of the total perceived cognitive workload, in five of total six individual cognitive workload categories (except for 'mental demand'), no significant differences among the tested approaches were observed (Table 10).

6.4. Limitations

The core of the current thesis was to conduct a wayfinding experiment. Due to the existence of certain technological and software limitations, the experiment was performed based on the 'Wizard of Oz' methodology. Although this method is widely used for answering research questions that would not be possible to be answered otherwise, it is important to take into account that there are limitations, mainly due to the errors and effects the use of the method may cause. The choice of this methodology was based on the fact that the experimenter had to control participant's device without being noticed.

The experimenter had to deliver a fully functioning routing application system to the participant, by updating the location of the participant in the routing map application, according to the real position of the participant in space. In order to do so, the experimenter had to observe the participant's device and trigger at the right moment the switching to the next image. This was not such a demanding task since, based on the design procedure followed for the development of the custom-made routing application, the user's current location in two successive images differed approximately 2.5 m. In other words, the experimenter had in mind that approximately after every three or four steps of the participant, the map had to be updated. Nevertheless, due to the fact that not every participant was walking in the same pace during the navigation task, the updating of the map was not ideal in every case.

Since the experiment was a real-world case study, limitations related to the indoor environment emerged. All the experiments took place within a period of time of one month. At that time, in the testing area (E and D floors in the HIL building, Campus Hönggerberg, ETH Zürich), the exhibition of postgraduate theses of architectural students took place. This fact, on one hand, limited the available space within the routing path, but on the other hand did not cause further problems to the participants during the wayfinding task. In addition, the time during the day the experiments took place was important, especially with regard to the total time that participants needed to complete the wayfinding task. All experiments were conducted from 10am to 6pm. Depending on the exact time during the day each experiment took place, the number of people interacting within the test area was quite different. This fact may have affected participants' navigation performance.

7. Conclusions and Outlook

The objective of the current work was twofold. On the one hand, the feasibility of the ILNM was investigated by combining it with indoor route maps, as well as by conducting a wayfinding experiment with human participants. On the other hand, of great significance was also the design and development of three cartographic visualization scenarios, two of which were based on ILNM implementation, while the third one was based on the benchmark approach for indoor navigation assistance (metric-based routing information).

The experiment's results clearly indicated the beneficial effect of ILNM implementation, in terms of participants' navigation performance. More specifically, the number of participants' errors observed during the testing of the two landmark-based approaches, was significantly lower compared to the one that occurred during the testing of the benchmark approach for indoor navigation assistance (metric-based). The inclusion of landmarks, not only within the routing instructions, but also as visualizations on the base-map itself, facilitated and improved the navigation performance of the participants.

Two of the main dimensions of participants' results analysis were concerning their user experience as well as their overall perceived cognitive workload during the interaction with the three different scenarios. Regarding participants' user experience, it was quite apparent that the metric- based approach was ranked as quite conservative, compared to the two landmark-based approaches. Furthermore, in terms of the user experience category of attractiveness, which is a clear indication of user's overall impression about a system, the two landmark-based approaches scored higher than the metric-based approach. However, in terms of cognitive workload, participants who were recruited for testing the metric-based approach needed to exert significantly lower amounts of mental effort during the wayfinding task compared to those who were testing the two landmark-based approaches. A possible explanation for this specific result could be the fact that landmark representations, especially those based on axonometric design approach, contained a high level of detail, something that was confusing for many participants. Due to this fact, many of them needed to exert higher mental effort in order to fully interpret each and every of the included landmark representations, something that can surely explain the aforementioned statistical result.

Summarizing, the current work can be further improved on the basis of the following proposals:

- Conduct of wayfinding experiment on larger scale: The design and conduct of a wayfinding experiment, for which more participants will be recruited. Also important would be the selection of a more diverse group of participants (similar number of male and female participants, different age groups). This will lead to more concrete and substantial results, mainly in terms of participants' user experience and cognitive workload analysis among the three different cartographic scenarios. Moreover, also beneficial in that context will also be the selection of a larger test area with more diverse sub-paths along the route, in terms of navigation difficulties.

- Design and conduct of a wayfinding experiment, in which experienced and non-experienced participants within the testing area will be included: Investigate possible differences in terms of navigation performance between these two different participants' categories.

- Examine other possible ways for landmarks visualization: Important in the further improvement of the current work would be the examination of other forms of representing the selected landmarks within the route maps. For instance, interesting would be the design of a cartographic visualization scenario with the inclusion of text-based or even sketch-based landmarks representations. Moreover, an important aspect must be the adoption of design approaches with the same level of detail complexity. This can lead, on the one hand, to landmarks visualizations that are semantically and visually equivalent, and on the other hand, to more solid conclusions regarding the evaluation of the cartographic representations of the landmarks.

- Evaluation of cartographic visualization approaches in different mobile screen sizes: The use of different mobile screen sizes in this experimental framework will help to evaluate in a more unbiased way landmark visualization approaches, which under the current conditions have been characterized as confusing, due to their high level of detail. More specifically, approaches such as those based on axonometric design philosophy, will be better and more objectively assessed.

- Research on formulating general guidelines for describing landmarks, based on the type of the building: One of the main limitations of the ILNM is the need of a group of experts to evaluate the suitability of the indoor spatial feature categories that may serve as landmarks. On the basis that there are similarities among the indoor spatial objects that exist in buildings of the same type (e.g., hospitals, universities), research on formulating general guidelines for objectively assessing these indoor spatial features, based on the type of the building in which they lie, will be beneficial for the further expansion of the algorithm. In that context, the need of candidate landmarks' assessment by a group of experts, will no longer be necessary.

Author Contributions: Conceptualization, N.B.; C.G.; Formal analysis, N.B.; Investigation, N.B.; Project administration, L.H.; Resources, L.H.; Supervision, C.G.; L.H.; Validation, C.G.; Writing—original draft, N.B.; Writing—review & editing, C.G.; L.H.

References

1. Hölscher, C.; Meilinger, T.; Vrachliotis, G.; Brösamle, M.; Knauff, M. Up the down staircase: Wayfinding strategies in multi-level buildings. *J. Environ. Psychol.* **2006**, *26*, 284–299. [CrossRef]
2. Raper, J.; Gartner, G.; Karimi, H.; Rizos, C. A critical evaluation of location based services and their potential. *J. Location Based Serv.* **2007**, *1*, 5–45. [CrossRef]
3. Kohtake, N.; Morimoto, S.; Kogure, S.; Manandhar, D. (Eds.) Indoor and outdoor seamless positioning using indoor messaging system and GPS. In Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN2011), Guimarães, Portugal, 21–23 September 2011.
4. Liu, J.; Chen, R.; Pei, L.; Guinness, R.; Kuusniemi, H. A hybrid smartphone indoor positioning solution for mobile LBS. *Sensors* **2012**, *12*, 17208–17233. [CrossRef] [PubMed]
5. Mendelson, E. System and Method for Providing Indoor Navigation and Special Local Base Service Application for Malls Stores Shopping Centers and Buildings Utilize RF Beacons. United States Patent US 8,866,673, 21 October 2014.

6.   Rehrl, K.; Häusler, E.; Leitinger, S. (Eds.) Comparing the effectiveness of GPS-enhanced voice guidance for pedestrians with metric-and landmark-based instruction sets. In *International Conference on Geographic Information Science*; Springer: Berlin/Heidelberg, Germany, 2010.

7.   Gkonos, C.; Giannopoulos, I.; Raubal, M. Maps, vibration or gaze? Comparison of novel navigation assistance in indoor and outdoor environments. *J. Loc. Based Serv.* **2017**, *11*, 29–49. [CrossRef]

8.   Hirtle, S.C.; Raubal, M. Many to many mobile maps. In *Cognitive and Linguistic Aspects of Geographic Space*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 141–157.

9.   Raubal, M.; Winter, S. (Eds.) Enriching wayfinding instructions with local landmarks. In *International Conference on Geographic Information Science*; Springer: Berlin/Heidelberg, Germany, 2002.

10.   Duckham, M.; Winter, S.; Robinson, M. Including landmarks in routing instructions. *J. Loc. Based Serv.* **2010**, *4*, 28–52. [CrossRef]

11.   Fellner, I.; Huang, H.; Gartner, G. "Turn Left after the WC, and Use the Lift to Go to the 2nd Floor"—Generation of Landmark-Based Route Instructions for Indoor Navigation. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 183. [CrossRef]

12.   Epstein, R.A.; Vass, L.K. Neural systems for landmark-based wayfinding in humans. *Phil. Trans. R. Soc. B* **2014**, *369*, 20120533. [CrossRef] [PubMed]

13.   Siegel, A.W.; White, S.H. The development of spatial representations of large-scale environments. Advances in child development and behavior. *Adv. Child Dev. Behav.* **1975**, *10*, 9–55. [PubMed]

14.   Kelley, J.F. (Ed.) An empirical methodology for writing user-friendly natural language computer applications. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, Boston, MA, USA, 12–15 December 1983.

15.   Laugwitz, B.; Held, T.; Schrepp, M. (Eds.) Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*; Springer: Berlin/Heidelberg, Germany, 2008.

16.   Schrepp, M.; Hinderks, A.; Thomaschewski, J. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *Int. J. Interact. Multimed. Artif. Intell.* **2017**, *4*, 103–108. [CrossRef]

17.   Schrepp, M.; Olschner, S.; Schubert, U.; User Experience Questionnaire (UEQ) Benchmark. Praxiserfahrungen zur Auswertung und Anwendung von UEQ-Erhebungen im Business-Umfeld. Tagungsband UP13. 2013. Available online: www.ueq-online.org (accessed on 26 March 2019).

18.   Hegarty, M.; Richardson, A.E.; Montello, D.R.; Lovelace, K.; Subbiah, I. Development of a self-report measure of environmental spatial ability. *Intelligence* **2002**, *30*, 425–447. [CrossRef]

19.   Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advances in psychology. *Adv. Psychol.* **1988**, *52*, 139–183.