Fighting Technology Dumb Down: Our Cognitive Capacity for Effortful AR Navigation Tools

James Wen[1], Agnes Deneka[2], William S. Helton[1],
Andreas Dünser[3], and Mark Billinghurst[1]

[1] University of Canterbury, Christchurch, New Zealand
[2] University of Twente, Enschede, The Netherlands
[3] CSIRO, Tasmania, Australia

Abstract. By overlaying virtual guidance information directly over the surrounding environment, Augmented Reality (AR) is seen as an easy alternative to maps for pedestrians navigating in unfamiliar urban environments. It is hypothesized, however, that easing navigation tasks would result in weaker cognitive maps, leaving users more vulnerable to becoming lost should their navigation device fail. We describe an outdoor navigation study that highlighted the gap between theoretical expectations and real world testing with navigation tools. We addressed the issues by creating a simulation system for testing navigation tools and report on the results of a study comparing AR with maps. We then extended the system to support simultaneous secondary tasks to assess relative workload. We present this as a way of objectively measuring relative cognitive effort expended on navigation tool use. Our findings are helpful in the design of mobile pedestrian navigation tools seeking to balance navigational efficiency with mental map formation.

Keywords: pedestrian navigation, augmented reality, maps, cognitive load, virtual environment, spatial knowledge acquisition.

1 Introduction

Rapid advances in mobile technologies coupled with the fact that many people find maps difficult to use [1] has led to the creation of a wealth of creative digital alternatives to traditional cartographic maps [7,8]. One promising technology is augmented reality (AR) where virtual information is overlaid directly upon the user's environment through a smartphone or digital eyewear [6,10]. It is assumed that users would complete navigation tasks far more efficiently with AR-based than with map-based interfaces in part because AR eliminates the lateral mental rotations as well as landmark feature associations required of map users [14]. It is hypothesized, however, that users of tools that make navigation tasks easier would do so at the expense of the user being spatially aware of their surroundings, which would leave them more vulnerable to getting lost should their navigation devices fail (e.g. due to battery depletion) [11,17]. We wish to assess the capacity of users to balance the acquisition of spatial knowledge (SK) while navigating with acceptable efficiency.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2014, LNCS 8512, pp. 525–536, 2014.

2 Background

The possible relationship between the ease-of-use of navigation tools and SK acquisition has been referred to as the “tradeoff hypothesis” where the improvement of one is at the expense of the other [11]. It has also been proposed that interfaces are becoming too easy to use and that more “effortful” interfaces need to be created so that the user will be forced to do work to help gain spatial awareness [5]. Indeed, it has been observed that users given random orientation quizzes during a navigation task performed significantly better in directional recall when compared to users not given such quizzes although it was noted that it may be unrealistic to expect users to accept having to answer random quizzes [12]. An attempt to integrate SK improvement features more tightly into the user interface by using landmark cues to encourage users to be more attentive to their surrounding environment, however, yielded no significant results [15].

Neither the notion of ease-of-use nor the nature of cognitive maps are easily quantified so it is challenging to test theories that relate the two. In our research, we seek to find objective and quantifiable measures that would provide insights for expended effort in using pedestrian navigation tools and SK acquisition for recalling routes traveled. We assume that AR is an easier interface to use than maps but hypothesize that maps would result in better SK, as shown in Figure 1. In this figure, the left half is considered *Easy* as less effort is required. This

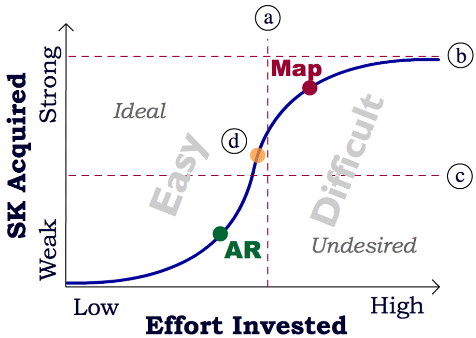

Fig. 1. The Spatial Knowledge - Effort curve

would be the desired half for interface usage. The upper half corresponds to stronger acquired SK, which would be the desired half for avoiding undesired dependencies on navigation technology. Maps would presumably reside in the upper right quadrant, being difficult to use but resulting in good SK acquired. AR, on the other hand, should be in the lower left quadrant: it is easy to use but results in weaker acquired SK.

Line ⓐ is calibrated by the tool designer and represents the maximum effort that could be reasonably imposed upon a user. To the right of this line, the interface may be considered to be too challenging to use. Line ⓑ refers to the maximum SK that can be acquired, such as a complete and thorough survey knowledge of the area. Line ⓒ corresponds to the desired minimum amount of SK the designer wishes for the user to gain. We propose a curve that passes through the AR and map positions and regard the portion of the line that passes through the *Ideal* upper-left quadrant as the interface we wish to design. Since the limits ⓐ and ⓒ are defined by the designer, the two can be adjusted so as to accommodate an ideal interface at ⓓ, which corresponds to a tool that balances effort invested with SK acquired.

In the remainder of this article, we report on our studies that sought to validate our assumption and test our hypothesis. We then describe a study measuring the capacity of users for accepting interfaces that attempt to increase user effort in order to improve SK so that an ideal interface for balancing navigation efficiency and acquired SK can be created.

3 Study 1: Performance Measurement

In order to support our assumption that AR-based pedestrian navigation tools are easier to use than map-based tools, our first study measured time-on-task performance and user perception of expended effort. A within-subjects design with three counter-balanced paths was employed on a university campus. Three smartphone interface conditions were tested: map, AR, and a combination mode that allowed the user to choose either the map or AR interface at any time (see Figure 2). Participants were asked to complete navigation tasks with a different

Fig. 2. The map (left) and AR (right) navigation interfaces

interface condition for each of the paths. At the end of each task, participants filled out a task load index (TLX) questionnaire. After completing all three paths, participants were asked to rank the interfaces in order of preference. We summarize our results and refer the reader to [2] for further details.

3.1 Results

A total of 22 participants (11 female, mean age = 31.0) completed the study. An ANOVA detected no significant differences in the time taken to navigate the

Table 1. Performance and TLX

	Time (sec)	TLX
	Mean (SD)	Mean (SD)
Map	919 (271)	39.58 (17.99)
AR	953 (344)	41.14 (17.04)
Combo	968 (275)	43.94 (15.69)

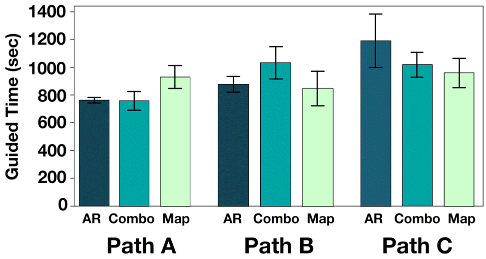

Fig. 3. Performance times for the paths

paths using the three interface conditions ($F(2, 38) = .25, p = .78$), as shown in Table 1 and Figure 3.

No significant differences were detected for the combined TLX workload scores between the interfaces ($F(2, 63) = .37, p = .69$). The ranking data showed that the AR interface was perceived to be significantly less useful than the Combination interface for completing the task ($\chi^2 = 10.30, df = 2, p < .01$). There was also a preference for the Map or Combination interfaces over the AR interface for use in everyday life ($\chi^2 = 6.90, df = 2, p < .05$).

3.2 Observations

Counter to our expectations, the AR interface did not offer more efficient traversal times. This, along with the results of a later independent study [13] led us to believe there may exist fundamental issues that were affecting how users perceive, use, and perform with AR-based pedestrian navigation tools. From observations, user feedback, and discussions, we suspected that outdoor tracking inaccuracies may be the primary factor causing this.

The issue of poor outdoor tracking is a known problem with considerable resources devoted to it. Since it will not likely be solved in the immediate future [9], we created a testbed within a virtual environment (VE) so that location information would be precisely defined at all times. By removing the tracking errors making AR difficult to use, we were able to focus our attention on defining the limits of how AR may affect navigational effectiveness and SK acquisition in an ideal environment.

4 Study 2: Simulating Perfect Tracking

We used the Unity3D game engine to build SPART (Simulator for Perfect AR Tracking) in order to evaluate user performance, perceived effort, and route recall in traversing a path through a desktop virtual environment. Using the mouse to change turn the direction of view and the w key to control forward movement, each participant was given navigation tasks that included a guided traversal followed by an unguided recall traversal of the same path. We employed a between-subjects design where each participant navigated two counter-balanced paths and was assigned one of three navigation interfaces: map (MP), map with You-are-Here marker (MY), and AR. These were activated by holding down the 1 key and displayed as shown in Figure 4. Our results are summarized here and we refer the reader to [16] for further details.

Fig. 4. The map (left) and AR (right) navigation interfaces

4.1 Results

A total of 71 participants (21 female, mean age = 19.8) successfully completed the study. There was a significant difference in guided navigation time depending on the tool being used (see Figure 5, left). MP users required the longest time and AR users needed significantly less time at the $p < .05$ level. MY users were between MP and AR users with no significant differences in performance when compared to either the MP or AR conditions.

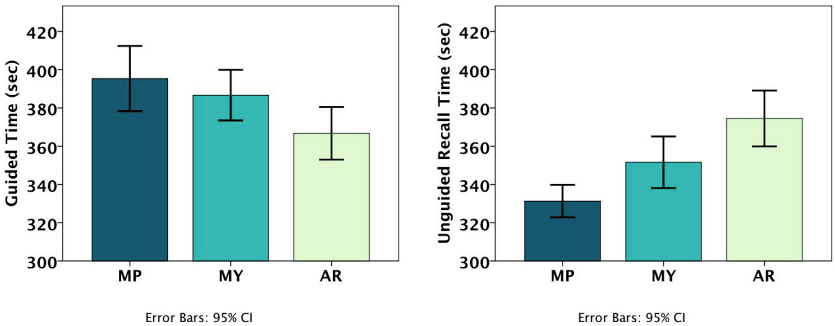

Fig. 5. Guided time (left) and unguided recall time (right)

Significant differences were also observed between interfaces for the unguided recall times (see Figure 5, right). MP users now took the least time while AR users took significantly more time at the $p < .001$ level. MY users were again between the MP and AR users; they were significantly slower than the MP users and significantly faster than the AR users, both at the $p < .05$ level.

A comparison of the unguided recall times and the initial guided traversal times is shown in Figure 6 (left). An independent t-test indicated significant differences in the average guided time and the average unguided recall time for MP users at the $p < .001$ level. Significant differences were also found for MY users in guided traversal times and unguided recall times at the $p < .001$ level. No significant differences were found for AR users between the average guided time and the average unguided time.

There was a significant difference in perceived workload between the interfaces (see Figure 6, right). A significant difference was observed between MP and AR, ($F(2, 66) = 6.46, p < .01$). MY yielded no significant differences with MP ($p = .20$) and AR ($p = .27$).

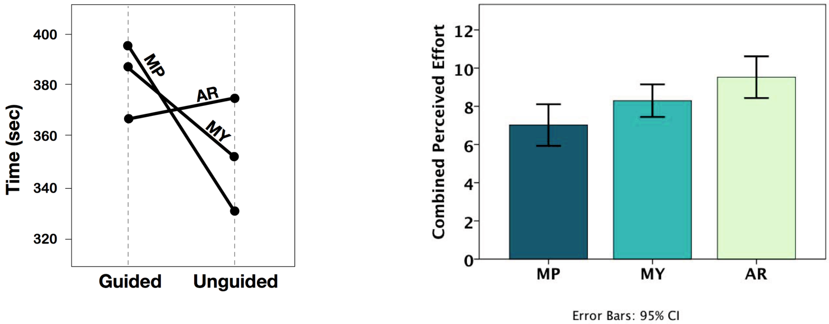

Fig. 6. Interaction (left) and TLX (right)

4.2 Observations

The results aligned with our assumptions. Map users took longer to navigate through the environment than AR users. The additional time was associated with periods where the navigation tool was active and was therefore considered to be the user interpreting the map. Given the simpler and more direct nature of AR, little time was required for interpretation. In this way, the simulation seemed consistent with the aim of AR-based navigation tools, which is to simplify the task of navigation resulting in shorter travel time. The fact that such behavior has not been observed in the real world suggests that AR is perhaps not yet ready to be used as an effective GPS-based navigation system.

The reversal of the traversal time relationship between the navigation interfaces for guided traversal times and unguided traversals for VE participants (see Figure 5) also agrees with expectations: the greater amount of time spent navigating through the environment correlated with more efficient time retracing the path at a later point. Referring back to Figure 1, this is consistent with the placement of maps in the upper half of the graph and AR in the lower half.

The results of the TLX indicated that AR users perceived greater effort was expended than map users, which is also consistent with expectations since the TLX was administered after the unguided recall navigation task, which we expected to be more difficult for AR users. The TLX is, however, a self-assessment survey and, as such, a subjective measure. To create the design parameters we seek, we would like to find a more objective measure which we describe next.

5 Study 3: Measuring Effort Capacity with Dual Tasks

We attempted to find an objective and quantitative measure that captures the ease users have with AR and map-based pedestrian navigation tools. To do this, we created a secondary task that served to distract the user from the primary task of navigation. By measuring user performance in this dual task environment, we gained insights into how the competition for cognitive resources affected both the primary and secondary tasks which, in turn, provided insights into the level of cognitive effort required.

We modified SPART to periodically display a word in the upper left hand corner of the screen. The words were taken from a standard list that has been used in studies investigating the impact of secondary tasks upon the performance of primary tasks, such as [3]. Participants were first asked to navigate through the virtual city with a navigation tool while trying to memorize as many of the words as possible. After reaching the end of the path, the participants were given two minutes to recall as many of the memorized words as possible. They were then asked to re-traverse the same path but without a navigation tool.

5.1 Results

A total of 47 participants (25 female, mean age=26.3) completed the study. Comparing interfaces, map users were significantly slower than AR users during the guided traversal ($t_{45} = 2.86, p < .05$) but, during the unguided recall traversal, map users were significantly faster ($t_{45} = -3.06, p < .05$). This is shown in Figure 7 (left).

Comparing guided and unguided recall modes, map users were significantly faster in the unguided mode than in the guided mode ($t_{470} = 3.29, p < .001$). AR users, however, did not exhibit any significant differences in time between guided and unguided recall navigation ($t_{50} = -1.79, p = .08$). These are shown in Figure 7 (right).

As a between-subjects study, we were able to compare the data from this study with Study 2. After culling out MY participants and participants who traversed Path B first, we retained 22 participants for comparison. Map users exhibited significant differences in travel times between the single-task and dual-task modes for both guided ($t_{31} = -1.75, p < .05$) and unguided recall travels ($t_{31} = -2.44, p < .05$) (see Figure 8).

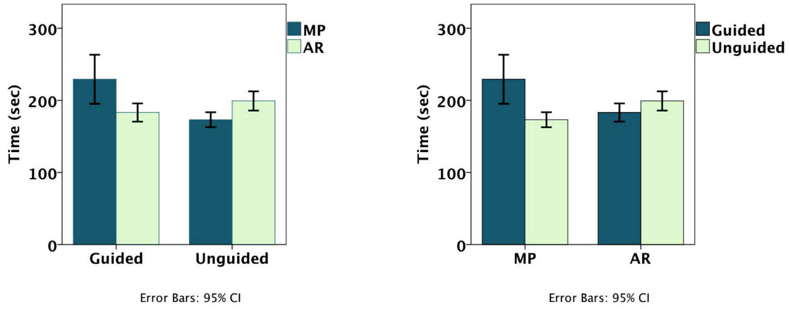

Fig. 7. Navigation performance comparisons for dual task study between interfaces and guidance modes

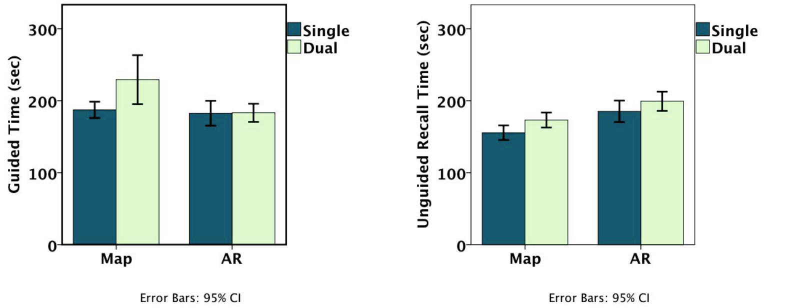

Fig. 8. Navigation performance comparisons between single and dual tasks

In contrast, AR users did not exhibit any significant differences in travel time between single-task and dual-task modes for either guided ($t_{36} = -.546, p = .59$) or unguided recall travel times ($t_{36} = -1.12, p = .27$). These are shown in the graphs of Figure 8. For the secondary task of word memorization, AR users recalled significantly more words than map users ($t_{45} = 2.08, p < .05$). A t-test conducted on the dual-task TLX results indicated that AR users perceived a significantly higher workload than map users ($t_{45} = -2.27, p < .05$).

5.2 Observations

The observation that guided performance suffered significantly for map users but not for AR users when switching from single to dual tasks suggest that the secondary task had a substantial impact on map users while it did not significantly affect AR users. A possible explanation for this is that map users had to split their cognitive resources between two tasks that were in competition. The AR condition, on the other hand, may have required so little cognitive effort that a secondary task would easily receive the attention it demanded without diminishing the performance of the primary navigation task. Users in the real world are often engaged in some secondary task—such as chatting with a friend—while finding their way in an urban environment, so the potential of AR-based navigation tools for effectively guiding such users without diminishing their secondary task is attractive.

While the secondary task slowed down map users significantly, it did not diminish the retention of a mental map for Map users since they were still significantly better than AR users in unguided travel time performance. This may suggest that the implicit retention of a mental map did not suffer as a consequence of having a secondary task. In other words, the secondary task increased the amount of time the user needed to accomplish the primary task but it did not seem to affect the mental map created.

With respect to the secondary task, AR users were able to recall significantly more words than map users without significantly increasing their guided traversal times from their single-task traversal. Map users, on the other hand, not only recalled significantly less words than AR users but also took significantly longer to follow the guided traversal when given a secondary task. As previously observed, this may show that AR users had greater spare cognitive capacity than map users to devote resources to the secondary task.

6 Discussion

Our interest in balancing navigation effectiveness with the formation of cognitive maps is challenged by a lack of objective quantitative metrics. Traditional approaches use distance and direction estimates as well as sketchmaps to build a model of cognitive maps but there are still fundamental arguments as to what actually constitute cognitive maps [4]. Instead of modeling a sophisticated concept, we chose to restrict our interest to a particular aspect of mental maps that is both more measurable and arguably of great interest: the retention of route knowledge. Our shift into a virtual setting proved advantageous since route knowledge can be tested with re-traversals but real world re-traversals of paths can be physically exhausting and not scalable on an experimental level. The difference in performance measures between the guided and subsequent unguided recall traversals gave us a basic measure for retained route knowledge.

We seek to balance SK acquisition with ease-of-use, which is another notion that is difficult to measure. Although the TLX survey is an effective measure, it is highly subjective. Our use of the dual task approach provides some basic insights in effort expended. Since our ultimate interest is in finding ways to introduce secondary tasks for improving SK, we do not need to transform effort expended into a potential ability for handling secondary tasks: the two are inversely related and so we simply interpret the results directly.

Given our focus on SK acquisition, the use of word memorization as a secondary task may lack substantial conflict with the faculties dedicated to spatial abilities. Future work should attempt to introduce secondary tasks that compete for resources more directly related to navigation.

Figure 9 illustrates how our measure can be used to guide interface design. Line ⓐ defines the minimal capacity deemed necessary for handling secondary tasks for improving SK. Line ⓑ defines the level of efficient performance. Box ⓒ defines the constraints within which pedestrian navigation tools would ideally occupy that could exploit the capacity of performing a secondary task for the sake of improving SK.

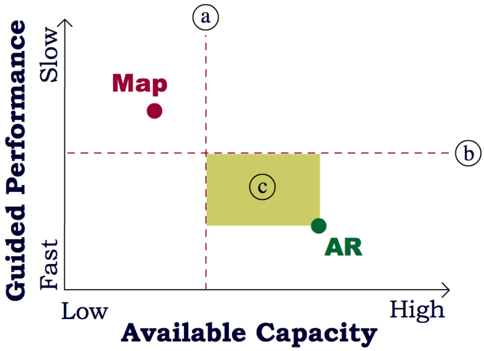

Fig. 9. Performance - Capacity Quadrants

Figure 10 illustrates how the relationship between time-on-task performance and acquisition of SK can be harnessed. The time saved by using an AR inter-

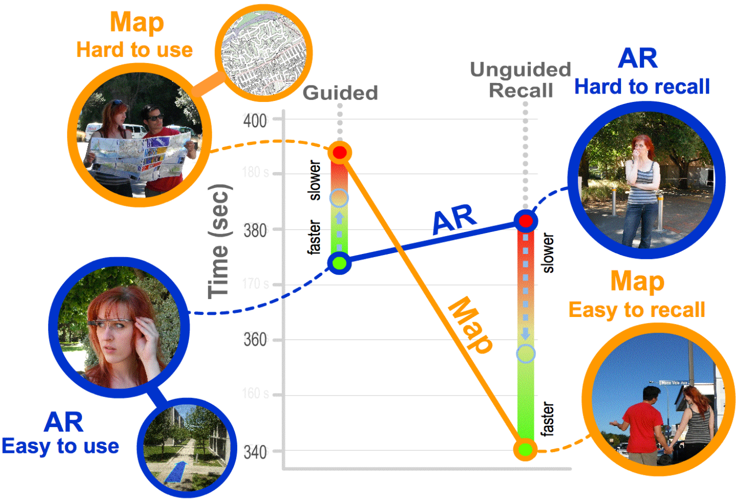

Fig. 10. Balancing navigation efficiency with SK acquisition

face for guidance (left side of graph) provides an opportunity to request more effort from the user in order to improve SK. For example, users can be asked to view a landmark briefly before receiving guidance directions. The dashed arrow indicates the time penalty caused by making AR more challenging. The benefit from the increasing the challenge is reflected in the unguided recall time (right side of graph) where the dashed arrow indicates the desired improvement in performance. The measures we presented here can help quantitatively calibrate the design of pedestrian navigation tools that employs the time saved using an AR-based navigation interface to offset the decrease in SK acquired without adversely affecting performance significantly.

7 Conclusion

There is a desire to create pedestrian navigation tools that provide efficient guidance information without sacrificing the formation of cognitive maps. To date, theoretical efforts have largely been qualitative and subjective. Experimental attempts have either yielded interfaces that would strengthen SK at the cost of usability or, they have been inconclusive. We are interested in finding an objective, quantitative measure for the factors that relate navigation effectiveness and a user's capacity to accept interfaces that require more effort for the sake of improving SK. We compared one of the newest technologies for pedestrian navigation with one of the oldest: AR with maps. We saw that AR is not quite ready for practical use although it may still, in principle, offer more efficient and better ease-of-use navigation guidance than maps. To explore this, we built SPART, a simulator that provided perfect location data for AR tracking. Our navigation testbed supported the assumption that AR-based navigation tools would be faster and easier to use than maps but offer weaker cognitive maps. By using a dual task approach, we were able to measure the capacity of users to take on secondary tasks that could be devoted to SK acquisition. We found that AR users had far greater capacity than map users to undertake a secondary task without penalizing the guided performance time. With this set of studies, further work in this area can now have a firm quantitative base upon which to design features that serve to balance the ease of navigation while improving the formation of SK.

References

1. Board, C.: Map Reading Tasks Appropriate in Experimental Studies in Cartographic Communication. Cartographica: The International Journal for Geographic Information and Geovisualization 15(1), 1–12 (1978)
2. Dünser, A., Billinghurst, M., Wen, J., Lehtinen, V., Nurminen, A.: Exploring the use of Handheld AR for Outdoor Navigation. Computers & Graphics 36, 1084–1095 (2012)
3. Green, A.L., Helton, W.S.: Dual-task performance during a climbing traverse. Experimental Brain Research 215(3-4), 307–313 (2011), http://www.ncbi.nlm.nih.gov/pubmed/22002634
4. Hirtle, S.C., Jonides, J.: Evidence of hierarchies in cognitive maps. Memory & Cognition 13(3), 208–217 (1985)

5. Huang, H., Schmidt, M., Gartner, G.: Spatial Knowledge Acquisition in the Context of GPS-Based Pedestrian Navigation. In: Zentai, L., Reyes Nunez, J. (eds.) Maps for the Future. Lecture Notes in Geoinformation and Cartography, vol. 5, pp. 127–137. Springer, Heidelberg (2012), http://www.springerlink.com/index/10.1007/978-3-642-19522-8
6. Liarokapis, F., Mountain, D., Papakonstantinou, S., Brujic-okretic, V., Raper, J.: Mixed Reality For Exploring Urban Environments. In: 1st International Conference on Computer Graphics Theory and Applications (2006)
7. May, A.J., Ross, T., Bayer, S.H., Tarkiainen, M.J.: Pedestrian navigation aids: information requirements and design implications. Personal and Ubiquitous Computing 7(6), 331–338 (2003), http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s00779-003-0248-5
8. Millonig, A., Schechtner, K.: Developing Landmark-Based Pedestrian-Navigation Systems. IEEE Transactions on Intelligent Transportation Systems 8(1), 43–49 (2007), http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4114345
9. Mulloni, A.: Enhancing Handheld Navigation Systems with Augmented Reality. In: Mobile HCI, pp. 5–8 (2011)
10. Mulloni, A., Seichter, H., Schmalstieg, D.: Handheld Augmented Reality Indoor Navigation with Activity-Based Instructions. In: Mobile HCI 2011 (2011)
11. Münzer, S., Zimmer, H.D., Baus, J.: Navigation assistance: a trade-off between wayfinding support and configural learning support. Journal of Experimental Psychology. Applied 18(1), 18–37 (2012), http://www.ncbi.nlm.nih.gov/pubmed/22141461
12. Parush, A., Ahuvia, S., Erev, I.: Degradation in spatial knowledge acquisition when using automatic navigation systems. In: Winter, S., Duckham, M., Kulik, L., Kuipers, B. (eds.) COSIT 2007. LNCS, vol. 4736, pp. 238–254. Springer, Heidelberg (2007)
13. Rehrl, K., Häusler, E., Steinmann, R., Leitinger, S., Bell, D., Weber, M.: Pedestrian Navigation with Augmented Reality, Voice and Digital Map: Results from a Field Study assessing Performance and User Experience. In: Proceedings of the 8th International Symposium on Location-based Services, pp. 3–20 (2011)
14. Seager, W., Fraser, D.S.: Comparing physical, automatic and manual map rotation for pedestrian navigation. In: CHI 2007, pp. 767–776 (2007), http://portal.acm.org/citation.cfm?doid=1240624.1240742
15. Waters, W., Winter, S.: A wayfinding aid to increase navigator independence. Journal of Spatial Information Science 3(3), 103–122 (2011), http://josis.org/index.php/josis/article/view/43
16. Wen, J., Helton, W.S., Billinghurst, M.: Classifying Users of Mobile Pedestrian Navigation Tools. In: Oz CHI, pp. 1–5 (2013)
17. Willis, K.S., Hölscher, C., Wilbertz, G., Li, C.: A comparison of spatial knowledge acquisition with maps and mobile maps. Computers Environment and Urban Systems 33(2), 100–110 (2009), http://linkinghub.elsevier.com/retrieve/pii/S0198971509000064