Strategies for Orientation: The Role of 3D Landmark Salience and Map Alignment

Clare Davies
Research Labs, Ordnance Survey,
Romsey Road, Southampton,
SO16 4GU, UK.
clare.davies@ordnancesurvey.co.uk

David Peebles
Department of Behavioural Sciences,
University of Huddersfield, Queensgate,
Huddersfield, HD1 3DH, UK.
D.Peebles@hud.ac.uk

Overview

- In contrast to the recent room-sized-space studies cited in the debate on orientation strategies, outdoor-space studies have long indicated a key role for landmarks rather than precise geometry.
- We conducted an experiment and eye movement study to investigate orienting without physical movement, matching a static scene view to a map to determine its direction.
- In our experiment we presented stimuli for which single-landmark matching was not the optimal strategy; the only unambiguous information available for matching was the map's 2D geometry which could also be abstracted from the scene.
- Despite this, most participants still chose a landmark-based strategy, demonstrated by performance deficits where such landmarks were present and by participants' eye movements and verbal protocols.
- We argue that the debate about orientation and the role of geometry should consider a broader range of evidence on spatial cognitive processes.

Strategies for orientation

Orienting oneself in an environment with a map is a common problem carried out in a variety of contexts. It requires one to match a direction in the visible scene with a direction on the map and is generally assumed to involve *mental rotation* to match the 2D and 3D representations. Two strategies for orienting are:

Geometric — study the geometry of the visible scene and derive a mental representation of the 2D shapes of the ground layout (as would be seen if viewed from above, i.e. from the map's perspective). It has been argued (e.g., Hermer & Spelke, 1994) that geometry is the default cognitive 'module' and the preferred default strategy for use in (physical) orientation.

Landmark — in situations where a salient cue or landmark can be identified in both the scene and the map, use the cue as an orientation indicator to match other items in the scene according to their position relative to it. Landmark matching may be more efficient in many circumstances, and has been shown to be used by adults, animals and sometimes young children (Cheng & Newcombe, 2005).

Studies of orientation (physical and otherwise) using outdoor scenes have shown a strong role for landmarks and topology, often overriding spatial geometry (e.g., Gunzelmann & Anderson, 2006; Pick, Heinrichs, Montello, Smith, & Sullivan, 1995; Warren, Rossano, & Wear, 1990). In the majority of these studies, however, salient landmark features in the scene also tended to be salient on the map.

Aims of the study

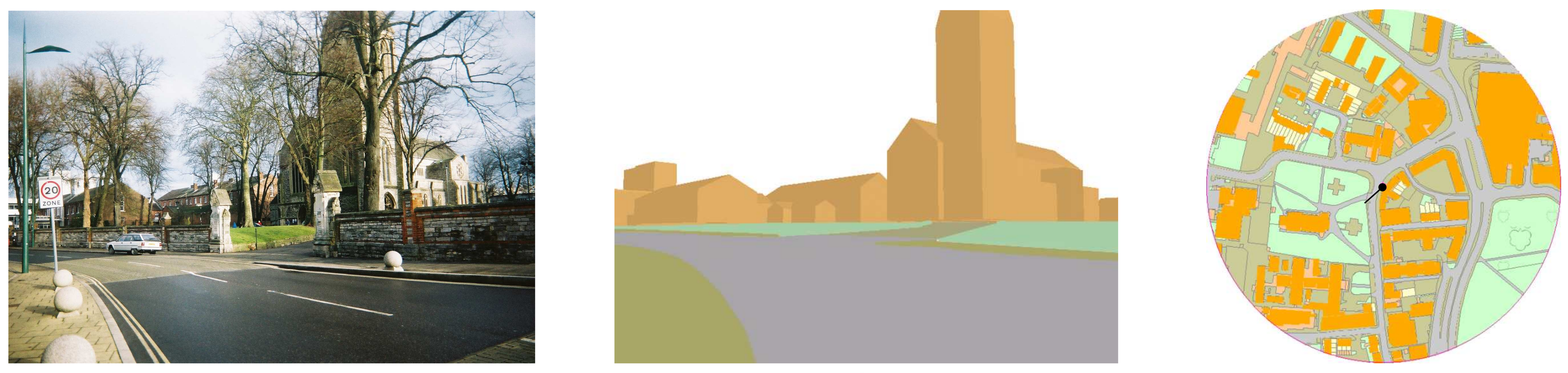
In this study we investigated these two strategies in a scenario where single-landmark matching would not be so easy, using geometrically irregular (European) urban spaces. We used pairs of scenes and maps taken from a 3D model of a UK city with only the 2D ground layout and the 3D building shapes being shown (e.g., Figures 1b and 1c and Figures 2b and 2c – the actual real-world streets corresponding to the scenes are shown in Figures 1a and 2a respectively). All irrelevant details scenes were removed from the scenes and the maps.

The scenes varied in terms of (a) the presence/absence of distinctive 3D landmarks and (b) the presence/absence of distinctive 2D layout cues. Only the 2D geometry could reliably solve the task however because the 2D outlines of distinctive (e.g., tall) 3D landmarks were ambiguous on the maps. Therefore, choosing a single item based on salient 3D cues (e.g., the large and distinctive 3D object in Figure 1b), and attempting to match it to the map was unlikely to be successful since its 2D geometry would probably not be sufficiently unambiguous on its own (but only when combined with other ground layout cues or relative object positions).

We hypothesised therefore that people would use the 2D geometry (e.g., roadside shape or relative object locations) to solve the orientation problem, rather than focusing on these visually salient but task-irrelevant 3D objects. If they were distracted by the latter, then performance would be worse for scenes such as Figure 1 than for those such as Figure 2 where heights and shapes were less variable, although 2D ground layout was often just as complex.

Experiment

Forty-nine students and members of staff from the University of Huddersfield took part in the experiment. An additional five participants carried out the experiment while having their eye movements and verbal protocols recorded.



(a) (b) (c)

Figure 1: Street location (a), scene (b) and corresponding map (c), stimulus 18. © Crown copyright 2007 Ordnance Survey.

The stimuli were 25 scenes and corresponding maps from various locations in the city of Southampton, UK. The scene images were generated using a buildings-only 3D model overlaid on OS MasterMap® Topography Layer and draped on an OS Land-Form PROFILE® terrain model to provide a realistic and accurate representation of height information. The maps were circular sections of OS MasterMap® Topography Layer at 1:1250 scale. A black dot in the centre of the map indicated the location of the observer. When the mouse cursor was moved over the map, a short black line of fixed length was drawn from the centre of the dot toward the tip of the cursor. This rotated around the dot as the mouse was moved around the map so that it always pointed toward the mouse cursor.

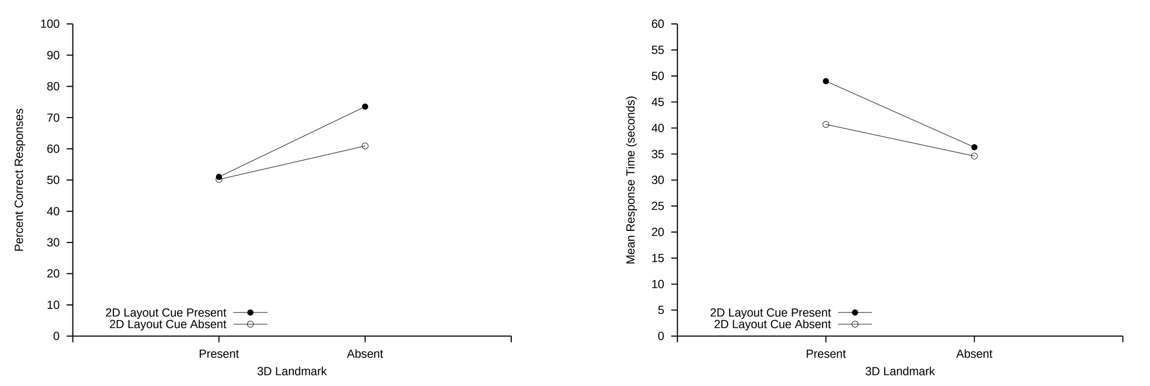
All participants saw the entire set of twenty-five stimuli (five practice trials and 20 experimental trials) in random order. On each trial, a scene and corresponding map were presented on the screen (e.g., the scene/map pairs shown in Figures 1b and 1c and Figures 2b and 2c). When the participant responded by clicking on the map, the latency and angular degree of the response were recorded. Participants in the eye movement and verbal protocol study were asked to talk through each trial as they attempted to solve the problem, in particular to say what they were looking at, how they were thinking through the problem, and why and how they were choosing a particular direction.

Results

Orientation strategies

Responses were scored as correct if the angle of the response line fell within 15 degrees of the true angle in either direction. The scenes were coded according to the presence or absence of 3D landmark and according to the presence of distinctive 2D ground layout information in the foreground of the scene.

The presence of salient 3D landmarks and distinctive 2D ground layout both had a significant effect on the accuracy of responses, ($F(1,48) = 40.35$, $p < .0001$, and $F(1,48) = 5.47$, $p < .05$ respectively). There was also a significant interaction between them, $F(1,48) = 5.26$, $p < .05$ (shown in Figure 3a). The directions of these effects showed that while presence of an obvious 2D cue was able to decrease error rates, this was only in the absence of a salient 3D cue which always greatly increased them.



(a) (b)

Figure 3: Percentage of correct responses (a) and mean response time (b) for stimuli categorised by the presence or absence of 2D and 3D cues

The analysis of response times, however, showed that both 3D and 2D cues seemed to slow participants down — much more so for 3D, $F(1,48) = 29.7$, $p < .0001$ than for 2D, $F(1,48) = 9.28$, $p < .005$. There was again a mild interaction, $F(1,48) = 4.37$, $p < .05$, (shown in Figure 3b) which indicated that the presence of both a 2D and a 3D cue had the most marked effect of all on response times; the presence of a 2D landmark made only a small difference except when a 3D landmark was also present.

This finding was independently confirmed by qualitative verbal protocol and eye movement analysis of the five additional participants. By far the most commonly reported feature used for solving the problem was 'buildings', and the eye movement patterns in the scenes with the most salient 3D landmarks (e.g., large skyscrapers or church steeples) tended to strongly focus around those landmarks.

Map alignment

Previous studies where a map is matched to a scene have tended to find a distinctive 'M shape' pattern in the effect of map alignment with observer position (e.g., Gunzelmann & Anderson, 2006; Hintzman, O'Dell, & Arndt, 1981). Performance typically is better not only at 0 degrees (where 'up' on the map exactly corresponds to the forward direction within the scene), but also at 90, 180 and 270 (i.e. -90) degrees. It seems that mental rotation to these cardinal directions is easier than with more oblique angles.

Comparing the RTs from our experiment with the M-shaped curve found in Gunzelmann and Anderson (2006), Experiment 1, however, it can be seen that although the M shape is partly visible in the RTs from our study, many scenes appear to violate it: indeed, the alignment angles for the three fastest scenes were -53, 76 and -17 degrees.

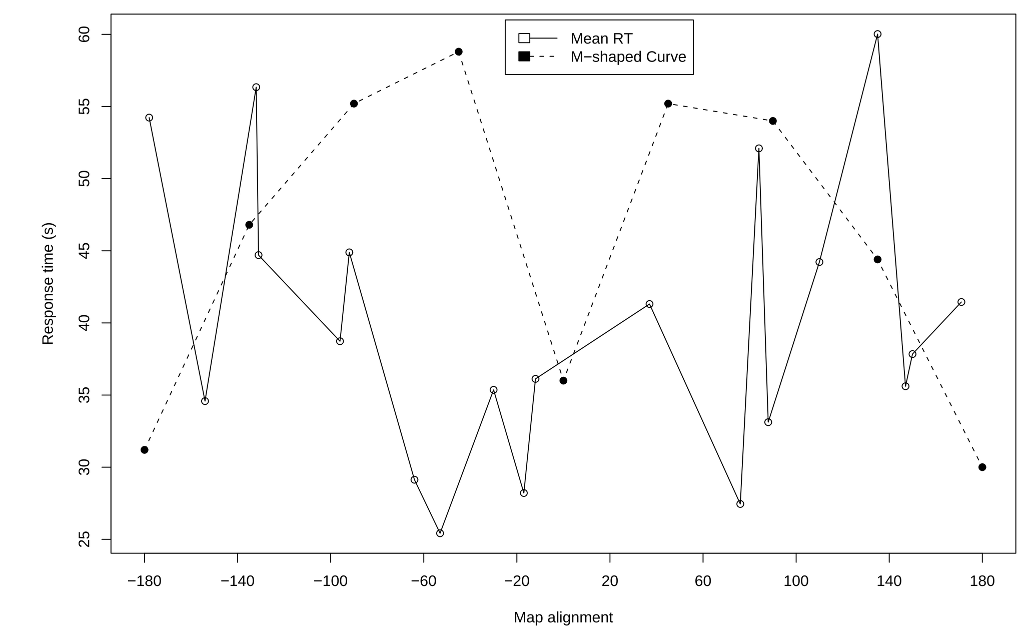


Figure 4: Response latencies plotted as a function of the map's alignment with observer position and compared with the M-shaped RT curve (rescaled) from Gunzelmann & Anderson (2006), Experiment 1.

Conclusions

- The results of this experiment further demonstrate people's tendency to match a single salient landmark between a 2D and 3D representation of a scene when orienting. Specifically, people selected landmarks with distinctive 3D (but not 2D) shapes despite the absence of 3D cues in the 2D map. This, along with the slower response times where a 2D ground layout cue was provided, implies that participants may find it difficult to abstract a 2D overhead layout from the 3D scene.
- These results can be seen to be at odds with the recent assumed preeminence of geometry as the primary source of orientation information for both humans and other animals (e.g., Hermer & Spelke, 1994; Cheng & Newcombe, 2005).
- The disruption of the usual 'M shape' effect of map alignment also indicates that map alignment alone (implying a strong role of mental rotation in the task) is not the only factor influencing orientation performance. The scenes which had unexpectedly good performance despite their alignment angle were apparently those where it was relatively easy to match an unambiguous cue to the map, regardless of its angle from the map's upward (north) direction.
- In this task (as in others such as way finding), simplified semi-topological encoding of a scene's key features may be more efficient than slow geometric transformations (e.g., mental rotation).
- To aid orientation with a map, it is likely that depicting appropriate (possibly 3D) landmarks on the map would improve performance. A broader range of spatial cognitive evidence, especially with more realistic scenes, should be considered in evaluating the claims made for geometry as a fundamental process.

References

Cheng, K., & Newcombe, N. (2005). Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychonomic Bulletin & Review*, *12*, 1–23.

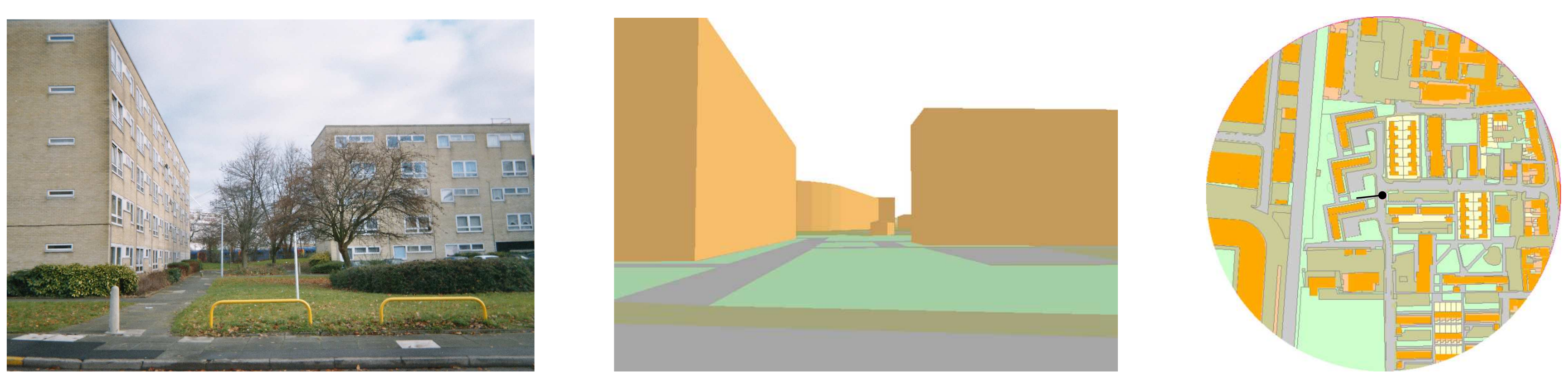
Gunzelmann, G., & Anderson, J. R. (2006). Location matters: Why target location impacts performance in orientation tasks. *Memory and Cognition*, *34*, 41–59.

Hermer, L., & Spelke, E. (1994). A geometric process for spatial representation in young children. *Nature*, *370*, 57–59.

Hintzman, D. L., O'Dell, C. S., & Arndt, D. R. (1981). Orientation in cognitive maps. *Cognitive Psychology*, *13*, 149–206.

Pick, H. L., Heinrichs, M. R., Montello, D. R., Smith, K., & Sullivan, C. N. (1995). Topographic map reading. In P. A. Hancock, J. Flach, J. Caird, & K. Vicente (Eds.), *Local applications of the ecological approach to human-machine systems* (Vol. 2, pp. 255–284). Hillsdale, NJ: Lawrence Erlbaum.

Warren, D. H., Rossano, M. J., & Wear, T. D. (1990). Perception of map-environment correspondence: The roles of features and alignment. *Ecological Psychology*, *2*, 131–150.



(a) (b) (c)

Figure 2: Street location (a), Scene (b) and corresponding map (c), stimulus 7. © Crown copyright 2007 Ordnance Survey.