Cognition 129 (2013) 24–30

Contents lists available at SciVerse ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Learning to navigate: Experience versus maps

Tobias Meilinger [a,b,*], Julia Frankenstein [a,c], Heinrich H. Bülthoff [a,d,*]

[a] Max-Planck-Institute for Biological Cybernetics, Spemannstr. 38, 72076 Tübingen, Germany
[b] Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan
[c] Center for Cognitive Science, University of Freiburg, Friedrichstr. 50, 79098 Freiburg, Germany
[d] Department of Brain and Cognitive Engineering, Korea University, Anam-dong 5ga, Seongbuk-gu, Seoul 136-713, Republic of Korea

ARTICLE INFO

Article history:
Received 15 November 2012
Revised 31 May 2013
Accepted 31 May 2013
Available online 30 June 2013

Keywords:
Spatial memory
Reference frame
Navigation
Route knowledge
Survey knowledge
Virtual environment

ABSTRACT

People use "route knowledge" to navigate to targets along familiar routes and "survey knowledge" to determine (by pointing, for example) a target's metric location. We show that both root in separate memories of the same environment: participants navigating through their home city relied on representations and reference frames different from those they used when doing a matched survey task. Tübingen residents recalled their way along a familiar route to a distant target while located in a photorealistic virtual 3D model of Tübingen, indicating their route decisions on a keyboard. Participants had previously done a survey task (pointing) using the same start points and targets. Errors and response latencies observed in route recall were completely unrelated to errors and latencies in pointing. This suggests participants employed different and independent representations for each task. Further, participants made fewer routing errors when asked to respond from a horizontal walking perspective rather than a constant aerial perspective. This suggests that instead of the single reference, north-up frame (similar to a conventional map) they used in the survey task, participants employed different, and most probably multiple, reference frames learned from "on the ground" navigating experience. The implication is that, within their everyday environment, people use map or navigation-based knowledge according to which best suits the task.

1. Introduction

People constantly find their way from one place to another – bedroom to bathroom; home to work – along familiar routes. Their wayfinding is guided at each decision point along the way by their underlying route knowledge (Golledge, 1999; Ishikawa & Montello, 2006; Mallot & Basten, 2009; Siegel & White, 1975; Thorndyke & Hayes-Roth, 1982; Trullier, Wiener, Berthoz, & Meyer, 1997; Wiener, Böchner, & Hölscher, 2009). By contrast, when people estimate distances and directions between mutually non-visible locations without necessarily knowing the connecting route they are informed by "survey knowledge". Route and survey knowledge of an area seem not to be tied together in a developmental sequence as often as suggested (Piaget, Inhelder, & Szeminska, 1960; Siegel & White, 1975); some navigators develop survey knowledge immediately, some over time, others never (Appleyard, 1970; Holding & Holding, 1989; Hölscher, Meilinger, Vrachliotis, Brösamle, & Knauff, 2006; Ishikawa & Montello, 2006; Moeser, 1988; Montello & Herbert, 1993). However, it remains unknown whether route and survey knowledge depend on different strategies operating on one representation (e.g., a mental map), or on different

* Corresponding authors. Address: Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan. Tel.: +81 3 5452 5249; fax: +81 3 5452 5249 (T. Meilinger). Address: Max-Planck-Institute for Biological Cybernetics, Spemannstr. 38, 72076 Tübingen, Germany. Tel.: +49 7071 601 601; fax: +49 7071 601 616 (H.H. Bülthoff).

E-mail addresses: tm@fennel.rcast.u-tokyo.ac.jp (T. Meilinger), heinrich.buelthoff@tuebingen.mpg.de (H.H. Bülthoff).

0010-0277/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2013.05.013

representations, and whether they use the same reference frame.[1]

To shed light on these questions we asked Tübingen residents to perform route and survey tasks with identical start and target locations, examining participant's knowledge of their home city acquired over years. Survey data were collected 1 week earlier as part of another study (Frankenstein, Mohler, Bülthoff, & Meilinger, 2012) and analyzed in conjunction with the route data. We hypothesized that if participants used a single representation for the two types of tasks their performance in both should be correlated. For example, a wrong turn on a route to a target location would correspond to a direction error when pointing to that location. Thus, more route errors should correspond to larger pointing errors. If they used different representations, however, no such correlations should be found. Related studies have not investigated this, having compared aggregated measures between participants, but not within participants' own performance (Appleyard, 1970; Hölscher, Büchner, Meilinger, & Strube, 2009; Ishikawa & Montello, 2006; Moeser, 1988; Montello & Herbert, 1993; Thorndyke & Hayes-Roth, 1982).

If people use different representations for route and survey tasks they might nevertheless use the same reference frames. For example, a photograph and a description of a scene are different representations, but they may use the same reference point and orientation. For survey knowledge, single reference frame representations have been described for learning simple environments from video or descriptions (Shelton & McNamara, 2004; Taylor & Tversky, 1992; Wilson, Tlauka, & Wildbur, 1999). In the population tested here, survey knowledge of one's city of residency is represented in a single, north-oriented reference frame likely acquired from maps (Frankenstein et al., 2012). When experiencing a complex environment by walking only, multiple local reference frames may be more important (Meilinger, Riecke, & Bülthoff, in press).

The reference frames underlying route knowledge have not been examined as much as those underlying survey knowledge. Theory states that route knowledge relies on multiple interconnected units (Mallot & Basten, 2009; Meilinger, 2008; Poucet, 1993; Trullier et al., 1997). These units (e.g., snapshots, local environments) serve to identify a location while their connections inform the navigator where to go next (e.g., a direction or street) or trigger a learned behavior. Thus route knowledge relies on multiple local reference frames. This has not, however, been demonstrated empirically until now.

To see whether route and survey knowledge reference frames differ, we varied the imagined perspective in which the route knowledge was recalled. Participants indicated their routing decisions (e.g., left, straight, right, etc.) both from an imagined horizontal, *walking* perspective and from a single imagined *aerial*, bird's eye or map perspective (Fig. 1). Spatial information is stored in a certain reference frame orientation, and accessing it from a different orientation usually yields inference costs such as errors or delays (McNamara, Sluzenski, & Rump, 2008). Otherwise it is classified as orientation-free. Performance measures may therefore reveal underlying reference frames. If the person is using a single reference frame it need be aligned once only with the aerial perspective during recall, but multiple times (i.e., after each turn) in the walking perspective (cf., paper map rotation during route navigation). Thus, we expect the person to perform better when doing the task from the aerial perspective. If participants use multiple local reference frames however, they should do better in the walking perspective (Meilinger, Franz, & Bülthoff, 2012) as the multiple frames would be identical with it and thus have no alignment costs.[2]

2. Methods

2.1. Participants

Twenty-three naïve participants (ten female), aged 18–50 years ($M = 28.5$; $SD = 7.7$) recruited from a subject database participated in exchange for monetary compensation after giving informed consent. They lived for at least 2 years in Tübingen ($M = 7.7$; $SD = 5.9$). All participants had performed the parallelized pointing task 1 week earlier (Frankenstein et al., 2012) and we reanalyzed these data. Two additional participants could not participate and additional two did not succeed in performing the task. The experiment was approved by the local ethics committee.

2.2. Materials

We used Virtual Tübingen, a highly realistic virtual model of Tübingen, Germany (see Fig. 2; http://virtual.tuebingen.mpg.de). Participants saw the model in horizontal perspective through a Kaiser SR80 head mounted display (HMD) while sitting on a swivel chair. Fog occluded adjacent intersections. We tracked head movements and rendered a stereo view of the virtual environment with a field of view of 63° (horizontal) × 53° (vertical) in real time. For further technical details see Frankenstein et al. (2012). We adjusted HMD fit and screen placement individually for every participant. The overall setup provided important depth cues such as stereo vision and motion parallax. Participants typed in route sequences with the arrow keys of a custom keyboard resting on their legs (see Fig. 2) and pointed in the identical setup using a custom made joystick.

[1] A reference frames is defined here as a reference location and orientation relative to which locations (and orientations) are represented.

[2] In walking vs. aerial testing, answers were given from a constant global reference frame or from reference frames changing after each turn. This difference was confounded with imagining a horizontal vs. a top down viewpoint. Any differences found might thus originate from answering from a constant vs. variable reference frames or from imagining a horizontal vs. top down viewpoint. To resolve the confound, participants could have been tested always from an imagined horizontal viewpoint, and indicated movement one time in the local street orientation as described and another time in a constant body orientation (e.g., always facing north), just without looking down from above. Unfortunately, this instruction was too complicated to understand. Therefore, we used the instruction confounding viewpoint and reference frame constancy. We address consequences for the interpretation of data in the discussion.