

# Verbally Assisting Virtual-Environment Tactile Maps: A Cross-Linguistic and Cross-Cultural Study

Kris Lohmann, Junlei Yu, Matthias Kerzel, Dangxiao Wang  
and Christopher Habel

**Abstract** The Verbally Assisting Virtual-Environment Tactile Maps (VAVE-TaM) approach proposes to increase the effectiveness of tactile maps by realizing an intelligent multi-modal tactile map system that generates assisting utterances that generates assisting utterances acquiring survey knowledge from virtual tactile maps. Two experiments in German conducted with blindfolded sighted people and with blind and visually impaired people show that both types of participants benefit from verbal assistance. In this paper, we report an experiment testing the adaptation of the German prototype to be useable by Chinese native speakers. This study shows that the VAVETaM system can be adapted to Chinese language with comparable small effort. The Chinese participants' achievement in acquiring survey knowledge is comparable to those of the participants in the German study. This supports the view that human processing of representationally multi-modal information is comparable between different cultures and languages.

---

K. Lohmann (✉) · J. Yu · M. Kerzel · C. Habel  
Department of Informatics, University of Hamburg, Hamburg, Germany  
e-mail: lohmann@informatik.uni-hamburg.de

J. Yu  
e-mail: jyu@informatik.uni-hamburg.de

M. Kerzel  
e-mail: kerzel@informatik.uni-hamburg.de

C. Habel  
e-mail: habel@informatik.uni-hamburg.de

D. Wang  
State Key Lab of Virtual Reality Technology and Systems, Beihang University,  
Beijing, China  
e-mail: hapticwang@buaa.edu.cn

**Keywords** Spatial knowledge acquisition · Virtual tactile map · Audio-tactile map · Multi-modal interface · Verbal assisting utterances · Virtual haptics

## 1 Introduction

Maps are important and highly successful means to acquire spatial knowledge about environments. Today, in the time of geographical information systems and the WWW, maps are more ubiquitous and easy to access than ever before. But, traditionally, printed maps as well as their modern counterparts, such as Google Map, are primarily designed and realized for sighted people. To provide access to adequate spatial information for blind and visually impaired people, tactile maps are used as substitutes [4]. Since reading tactile maps can be complicated and time consuming, tactile maps accessible via virtual haptic environments have been developed and verified to be usable by people with special needs [5]. Such maps can be generated on the basis of data available online [7].

To increase the effectiveness of spatial knowledge acquisition from tactile maps in virtual haptic environments, we developed the Verbally Assisting Virtual-Environment Tactile Map (VAVETaM) system, which combines spatial language and virtual variants of tactile maps [6, 9, 11]. The virtual variants of tactile maps can be perceived using the Sensable Phantom Omni haptic device ([www.sensable.com](http://www.sensable.com)). This device is placed on the desktop and allows a map user to explore three-dimensional virtual objects via force-feedback. In our case, the virtual objects are virtual variants of tactile maps generated on the basis of 3D models similar to those used in computer games and visual virtual environments. Objects on the maps are displayed as indentations and groves on a planar surface. To interact with these maps, the user holds and moves a pen-like handle attached to a small robotic arm mounted on the base of the device. The handle can be moved in all three dimensions. Sensors in the device measure the position of the handle and servomotors generate appropriate force-feedback, so that perceiving the virtual tactile map is enabled. The interaction can be pictured approximately like the exploration of a real object with a pen.

Empirical research conducted with assisting utterances in German has shown that blindfolded sighted people benefit from verbal assistance only with respect to their knowledge about landmarks (such as buildings), but not in their knowledge about tracks (such as streets) [10]. In contrast, blind and visually impaired people benefit in both types of spatial knowledge [8].

In this paper, we report on the potential for adapting the first, German-oriented, VAVETaM prototype to be useable by Chinese native speakers. Chinese instead of German assisting utterances were given and user experiments were conducted in China with Chinese native speakers. The results of the study, that has a cross-cultural and cross-linguistic perspective, have important implications. They show that the VAVETaM system for non-visual map access can be adapted to Chinese

language with comparable small effort and still be comparable efficient—besides all syntactic and semantic differences between the languages. Furthermore, they show that human processing of representationally multi-modal information (i.e., information that combines different representational systems, e.g., maps and language, see [1] for a discussion of representational modalities) is comparable between different cultures and languages.

## 2 Method

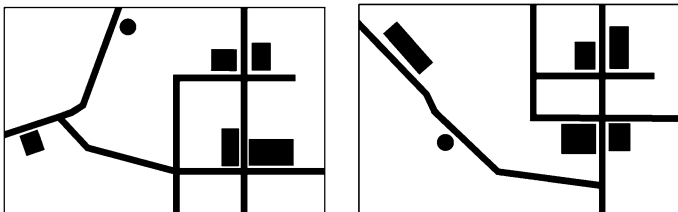
### 2.1 Participants

The data of 24 participants was evaluated (17 males, mean age: 25.4 years, SD: 2.7 years). All participants were compensated on a monetary basis. They were naïve about the purpose of the experiment. All participants gave written informed consent and reported to speak Chinese on a native-speaker level. All participants used their self-reported primary hand for the exploration of the virtual tactile map. Only one participant reported to be left-handed.

### 2.2 Materials and Procedure

*Maps.* The virtual tactile maps were explored using the Phantom haptic device. The material was similar to the experiment that was previously performed with German assistance [10], in the following referred to as the German-assistance experiment. The maps used were the same as those used in the German-assistance experiment. The maps included five tracks (track is a more general term for street-like structures) and six potential landmarks (one tree and five buildings). See Fig. 1 for a visualization of the maps used.

*Assisting Utterances.* The assistances were recorded before the study. They were given in Chinese, spoken by a 24-year-old male native Chinese speaker. The participants heard the utterances via headphones.



**Fig. 1** Visualization of the two maps (map 1 to the *left* and map 2 to the *right*) used in the experiment

Two sets of typical Chinese names that could be used for both maps were created (the number of streets and potential landmarks is identical on both maps). The set of assisting utterances that was used is a translated version of those used in the German study. The assisting utterances are grouped by message classes. A message class is defined by the information that an utterance of a class expresses. In the context of the present paper, the identification message class is important. An utterance of the identification message class informs the user of the virtual tactile map about the identity of the map object that is explored. Usually, this is done by stating the proper name of the map object in combination with a demonstrative. If an object does not have a proper name, e.g., an intersection, it can often be identified by referring to related other objects that carry proper names. For the intersection example, these are the names of the streets that form the intersection (see example (1d) below). A detailed discussion of the other message classes is out of the scope of this paper (see [9], for a detailed discussion of message classes). In the following, the assisting utterances that were included in the two different assistance conditions are discussed.

In the *identification-assistance condition* that provided the baseline for the described study, only utterances of the identification message class were included. Example (1a) is an English counterpart<sup>1</sup> of an assisting utterance that was given when the map object representing the street named ‘Main Street’ was explored. Examples (1b) and (1c) are translations of identification messages for the frame of the map and the building called ‘JFK High School’. In the present study, in addition to streets and landmarks that carry a proper name, intersections and dead ends were identified. For example, crossings were identified by utterances such as (1d) and dead ends were identified by utterance such as (1e).

(1a) This is Main Street.

(1b) This is the left map frame.

(1c) This is JFK High School.

(1d) This is the crossing Main Street Madison Avenue.

(1e) This is the dead end that forms the right end of Madison Avenue.

Additional information was given in the *extended-assistance condition*. (2a)–(2f) are examples for English counterparts of assisting utterances given in this condition when track objects were explored. (3a) and (3b) are translations of assisting utterances stated when a potential-landmark object was explored.

The extended-assistance condition included verbal information about the extend of tracks, i.e., what determines the end of a track was uttered [see example (2b)]. The set of utterances included information about the intersections a track had and spatial relations with other tracks and landmarks [see examples (2a), (2c),

---

<sup>1</sup> See the Appendix for examples of German and Chinese items used in the language specific experiments. The English counterparts in the paper are used to present the approach and the empirical studies independent from the linguistic characteristics of Chinese and German.

and (2d)]. Furthermore, for parts of tracks that were close to landmarks or between landmarks, assisting utterances were given that stated this relation [see (2f)].

For potential landmarks, the set of utterances in the extended-assistance condition included utterances that stated the relation to other map objects [see (3a)] and, if appropriate, the global location in the map [see (3b)].

(2a) Main Street is parallel to Hill Street.

(2b) Main Street ends to the left at an intersection with Church Street, and to the right in a dead end.

(2c) Below Main Street, there are JFK High School and St. Mary Church.

(2d) Main Street intersects with Madison Avenue.

(2e) This is the intersection between Main Street and Madison Avenue.

(2f) Now, you are between JFK High School and St. Mary Church.

(3a) JFK High School is located below Main Street.

(3b) St. Mary Church is located in the upper part of the map.

The experimenter started the playback of the utterances using custom-built software developed for that purpose. The experimenter saw the hand movement of the map user as a moving red dot on a visualization of the map. The experimenter started an utterance when the participant explored the corresponding map object. Color-indication was used to state which button belonged to which object on the map.

Earlier empirical studies indicated that if a map object is explored, giving information of the identification message class should precede any other information. Hence, the experimenter started an utterance of the identification class prior to any other assistance in the extended-assistance condition. The other utterances for that object were given in the order that the experimenter found most appropriate. However, we avoided unnatural repetitions of utterances. To ease this, the buttons for utterances that were started once were marked in the interface.

*Sketch-Mapping Task.* Following the exploration with one of the assistance conditions described, the spatial knowledge that the participants gained by the exploration of the map was tested. First, participants were asked to sketch the map they learned as precisely as possible. They were instructed to include all they remember. For a discussion of the reliability of sketch maps as testing methods for spatial knowledge see [2]. The frame defining the dimensions of the map was printed on the paper for sketching.

All 48 sketches were evaluated by two independent raters. The rating was performed in two respects: (1) to what degree does the sketch resemble the original map concerning the course of roads, their parallelism, and the junctions they have and (2) to what degree are potential landmarks represented at the correct position? The rating was performed on a 5-point Likert-type scale. A rating of 1 is associated with 'does not reflect the original' and 5 is associated with 'reflects the original precisely'.

*Configuration-Questions Task.* After participants completed the sketch map, they were asked to answer to questions about configurations of map objects. The

full test consisted of 20 questions. Where it was possible, those questions asked for object configurations between objects that were not explicitly stated in the pre-recorded assisting utterances. The experimenter stated the questions in a predetermined individual random order. 10 questions involved spatial relations including potential landmarks and 10 were only about the configuration of and relations between tracks. The answering options were 'yes', 'do not know', and 'no'. 10 questions were answered correctly with a 'no' and 10 questions with a 'yes'. A correct 'yes' and correct 'no' were evaluated as correct answers. A wrong 'yes', a wrong 'no', and 'do not know' were evaluated as wrong answers.

In the experimental procedure, each set of names occurred with each map. Therefore, 4 sets with 20 configuration questions were developed. The questions of those sets were matched with each other. For example, question 1 always asked for the relation of two landmarks that were relatively distant to each other and was always correctly answered with 'yes' and question 17 always asked if two streets are parallel and was always correctly answered with 'no'. Two sets of questions were created for each map. This resulted from the use of two different name sets, as described above. The individual questions for each name set always asked for a spatially equivalent fact. For example, in map 1 for name set 1 question 1 was 'Is the oak tree left of JFK High School?' (English names are used for illustrative purposes). In name set 2, the map object that was the 'oak' in name set 1 was called 'beech'. The map object that was called 'JFK High School' in name set 1 was called 'St. Patrick Church'. Consequently, question 1 for the map 1 with name set 2 was: 'Is the beech tree left of St. Patrick Church?'

*Procedure.* The experiment was performed as a repeated-measures experiment with assistance condition as within-subject variable. This design was chosen to overcome the problem of individual differences in the ability to understand virtual tactile maps resulting from differences in spatial abilities (see [13], for a discussion of individual differences in spatial abilities).

To avoid carry-over effects, two different maps and name sets were used for the different experimental conditions. To control for order effects that bias the main research question, the experiment was fully counterbalanced concerning the order of assistance conditions, maps, and name sets. Accordingly, the second learning procedure was performed with the map, the assistance condition, and the name set that were not used in the first condition. Consequently, the extended-assistance condition occurred as often as first learning condition as the identification-assistance condition. Furthermore, map 1 and name set 1 occurred as often as map 2 and name set 2.

To ensure that the participants understood the haptic interaction with the multi-modal system, they were trained in different aspects of it. To conclude the training, participants were tested for their ability to identify the shape of objects (such as a triangle and a square) and to follow a complex track structure without leaving it while receiving and following assistance given by the experimenter. The training procedure including the training test took 30–60 min.

After a short break, participants read written instructions about the time they had for exploration and the tasks they had to solve after map learning. The

information about the tasks was included to minimize order effects resulting from knowledge of the tasks from the first condition. For the same purpose, an example map similar to the ones used in the experimental conditions was printed on the instructions. Participants were instructed to read the map in a way that they would find the route from each potential landmark to each other.

Each exploration was limited to 8 min. Participants were informed when they had 2 min left. After the first exploration, first the relation-questions task was performed, then participants created the sketch map. A 10-min break was made before the participants performed the second exploration under the same time constraint as the first exploration.

## 3 Results

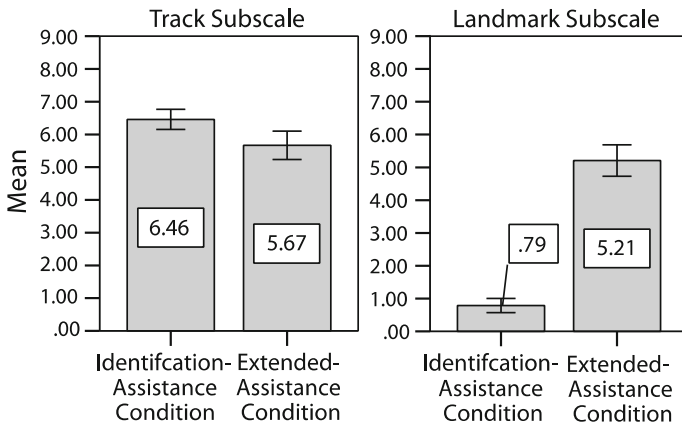
### 3.1 Configuration-Questions Task

Taking all questions into account, the average result showed that participants were able to answer significantly more questions correctly ( $t(23) = 5.03$ ,  $p < 0.001$ ) when they learned the map under the extended-assistance condition and the effect is large ( $r = 0.72$ ) (according to [3]).

An item reliability analysis was performed to control for reliability of the scales. For reasons that are explained in the remainder of this section, questions only about tracks and questions involving potential landmarks were regarded as two subscales and the reliability analysis was performed independently for each of the subscales. Two questions were excluded from further analysis as they did not increase overall reliability of the subscales. For the following analysis of the data, we assumed that the name set and the maps do not have an effect on the items (no significant effects of the map or the name set used were found in the analysis of the remaining 18 questions).

In the German-assistance experiment, the assistance condition had a strong effect on some, but not on all of the questions. Whereas the answers to the landmark-related questions showed large differences between the conditions, the answers to the questions that only involved tracks did not. Screening the dataset of the Chinese-assistance experiment, this finding could be replicated.

For the two resulting types of knowledge, two subscales were constructed and tested for reliability by calculating Cronbach's alpha. Since the questions involving landmarks differed from those involving only tracks, these subscales are named the landmark subscale and the track subscale. The landmark subscale with the remaining nine items provides a good reliability ( $\alpha = 0.848$ ). The track subscale with the remaining nine items provides an acceptable reliability ( $\alpha = 0.704$ ). See [10] for an English translations of the questions (of the set presented there, question 8 was excluded in the current experiment to improve reliability). Figure 2 shows the mean number of correct answers to the configuration questions after the learning conditions and for the two subscales. The figure indicates that extended



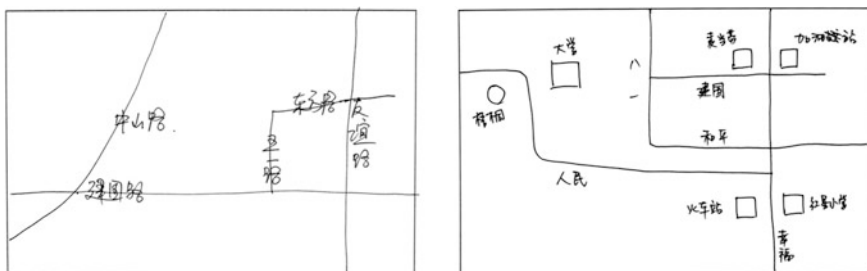
**Fig. 2** Mean number of correct answers in the track subscale and in the landmark subscale of the configuration-question task. Error bars represent the standard error of the mean

assistance leads to a larger amount of correctly localized landmarks and the knowledge about the track structure is comparable between learning conditions.

Separate mixed-design ANOVAs were calculated for each subscale. In both subscales, there were no significant main effects of gender, order of conditions, order of the maps used, or the order of the set of names ( $p > 0.05$ ). Therefore, these variables were stepwise excluded from the model for further analysis. Effects of handedness was not evaluated as only one left-handed person took part in the experiment.

In the track subscale, there was a non-significant trend that participants performed worse in the extended-assistance condition than in the identification-assistance condition ( $F(1,23) = 3.61, p = 0.07, r = 0.37$ ).

In contrast, there was a highly significant and large effect of the assistance condition on the amount of correct answers in the landmark subscale ( $F(1,23) = 69.10, p < 0.001, r = 0.87$ ). Participants were able to answer more questions correctly after having learned the map in the extended-assistance condition (Fig. 2).



**Fig. 3** Example of a participant’s sketch of map 1 after learning the map in the identification-assistance condition and example of a participants sketch of map 2 after learning the map in the extended-assistance condition



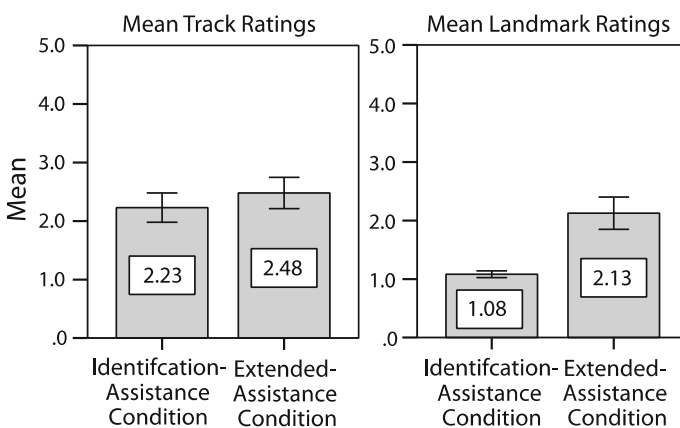
### 3.2 Sketch-Mapping Task

Two native Chinese raters evaluated the sketches (for examples, see Fig. 3). The intraclass-correlation coefficient was calculated to test the agreement between the raters. The agreements of the track ratings ( $ICC(3,1) = 0.81$ ) and of the potential-landmark ratings ( $ICC(3,1) = 0.95$ ) were calculated separately [12]. Both agreements are good. The values of the raters were averaged before the analysis. Figure 4 shows the mean ratings of the sketches. The values suggest that the sketches reflected knowledge about potential landmarks better when the map was learned under the extended-assistance condition.

Separate mixed-design ANOVAs were calculated for the landmark-configuration ratings and the track-structure ratings. In both ratings, there were no significant main effects of gender, handedness, the order of the maps used, or the order of the conditions. Therefore, these variables were stepwise excluded from the model for further analysis.

Concerning how well the track structure was reflected, no significant effect of the set of names was present in the data. Therefore, this between-subjects variable was excluded from the model as well. No significant effects of assistance condition were found ( $F(1,23) = 0.84, p > 0.05, r = 0.19$ ).

In contrast, the mean rating of how well the landmarks are represented in the sketches showed a significant main effect of the set of names used in the extended condition. Participants who performed the experiment in a setting in which name set 2 was used in the extended condition performed significantly better than participants which learned the map in the extended condition with the name set 1 ( $F(1,22) = 4.82, p < 0.05, r = 0.42$ ). Additionally, an interaction effect of assistance condition and name set was present in the data. Again, participants who performed the experiment in a setting in which the name set 2 was used in the



**Fig. 4** Mean ratings of the track structure and the landmark configuration on the sketches. Error bars represent the standard error of the mean

extended condition performed significantly better than participants which learned the map in the extended condition with the first name set ( $F(1,22) = 5.24$ ,  $p < 0.05$ ,  $r = 0.44$ ).

Most important for the purpose of the study, the ratings for how well potential landmarks were reflected in the sketches were significantly affected by the assistance condition ( $F(1,22) = 16.71$ ,  $p < 0.001$ ,  $r = 0.65$ ).

## 4 Conclusion

The study reported follows a study previously conducted in the German language. Both studies were performed with blindfolded participants. The reported study conducted in the Chinese language confirmed the finding that knowledge gained using the VAVETaM system consists of two subtypes: (1) knowledge about the structure of tracks and (2) knowledge about the configuration of potential landmarks. A clear effect of extended assistance is found in this study: Participants scored higher in all tests assessing the knowledge about potential landmarks. In contrast to the clear effect in the knowledge about potential landmarks, no test showed a significant effect of assistance condition on track knowledge. Rather, there is a non-significant trend in the data of the configuration questions that participants scored lower in questions of the track subscale. However, overall, participants benefit from extended assistance.

The study reported was performed in the framework of the development of the VAVETaM system, which is intended to help blind and visually impaired people. The results indicate that the system could be adapted to the Chinese language: Generally, the results are in line with the results obtained in the experiment conducted in Germany. This also indicates that the processing of representationally multi-modal information is to some extent culture and language independent.

## A.1 5 Appendix

**Table A.1** Name sets used in the German-assistance experiment [10] and corresponding name sets used in the Chinese-assistance experiment

	German names and English translation		Chinese names and English translation	
<i>Name set 1</i>	Poststraße	Post street	人民路	Renmin road
	Humboldtstraße	Humboldt street	解放路	Jiefang road
	Eiche	Oak	梧桐树	Parasol tree
	Hauptbahnhof	Central railway station	火车站	Railway station
	Christus-Kirche	Christ church	加油站	Gas station

(continued)

**Table A.1** (continued)

German names and English translation		Chinese names and English translation	
	Berthold-Brecht-Schule	Berthold-Brecht-school	红星小学 Hongxing primary school
<i>Name set 2</i>	Dorfstraße	Village street	东方路 Dongfang road
	Bergstraße	Hill street	建国路 Jianguo road
	Buche	Beech	柳树 Willow
	Rathaus	Town hall	法院 Court of justice
	Gedächtnis-Kirche	Memorial church	银行 Bank
	Anne-Frank-Schule	Anne-Frank-school	科技小学 Keji primary school

## References

- Bernsen NO (1994) Foundations of multimodal representations: a taxonomy of representational modalities. *Interact Comput* 6(4):347–371
- Blades M (1990) The reliability of data collected from sketch maps. *J Environ Psychol* 10 (4):327–339
- Cohen J (1992) A power primer. *Psychol Bull* 112(1):155–159
- Espinosa MA, Ungar S, Ochaita E, Blades M, Spencer C (1998) Comparing methods for introducing blind and visually impaired people to unfamiliar urban environments. *J Environ Psychol* 18:277–287
- De Felice F, Renna F, Attolico G, Distanto A (2007) A haptic/acoustic application to allow blind the access to spatial information. In: *World haptics conference*, pp 310–315
- Habel C, Kerzel M, Lohmann K (2010) Verbal assistance in tactile-map explorations: a case for visual representations and reasoning. In: *Proceedings of AAAI workshop on visual representations and reasoning 2010*, Menlo Park, CA, USA
- Kaklanis N, Votis K, Moschonas P, Tzovaras D (2011) HapticRiaMaps: towards interactive exploration of web world maps for the visually impaired. In: *Proceedings of the international cross-disciplinary conference on web accessibility*, Hyderabad, India
- Lohmann K (2013) Verbally assisting virtual tactile maps. An interface for visually impaired people. Dissertation, Universität Hamburg
- Lohmann K, Eschenbach C, Habel C (2011) Linking spatial haptic perception to linguistic representations: assisting utterances for tactile-map explorations. In: Egenhofer M, Giudice N, Moratz R, Worboys M (eds) *Spatial information theory*. Springer, Berlin, pp 328–349
- Lohmann K, Habel C (2012) Extended verbal assistance facilitates knowledge acquisition of virtual tactile maps. In: Stachniss C, Schill K, Uttal DH (eds) *Spatial cognition 2012*. Springer, Berlin, pp 299–318
- Lohmann K, Kerzel M, Habel C (2012) Verbally assisted virtual-environment tactile maps: a prototype system. In: Graf C, Giudice NA, Schmid F (eds) *Proceedings of spatial knowledge acquisition with limited information displays 2012*, Seeon, Germany
- Shrout PE, Fleiss JL (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 86(2):420–428
- Wen W, Ishikawa T, Sato T (2011) Working memory in spatial knowledge acquisition: differences in encoding processes and sense of direction. *Appl Cogn Psychol* 25:654–662