Ambiguity in Acquiring Spatial Representation from Descriptions Compared to Depictions: The Role of Spatial Orientation

Holly A. Taylor¹, David H. Uttal², Joan Fisher², and Marshall Mazepa²

¹ Tufts University, Department of Psychology, 423 Redwood Street, Medford, Massachusetts, USA 02155
htaylor@tufts.edu

² Northwestern University, Department of Psychology, 2920 Sheridan Road, Evanston, IL 60201-2970
duttal, joanfisher, m-mazepa@northwestern.edu

Abstract. Adults can make judgments about multiple spatial relations based on information gained from different kinds of input, including maps, descriptions, and through navigation [1]. However, factors such as spatial orientation influence performance. We investigated spatial orientation effects on learning from different media. In Experiment 1, participants learned a house from a map or a description. They then judged surrounding locations while imagining being in each room and they reconstructed the house. Participants who learned from a description performed worse on both tasks. Errors suggested they interpreted the term "in front" differently than intended [2]. Experiment 2 tested this hypothesis by examining two factors influencing interpretation of "in front", specific interpretation instructions and orientation information. The orientation information influenced performance more than the explicit interpretation of "in front." Taken together, the results indicate multiple influences on the spatial reference frame participants use to interpret spatial terms.

Key words: spatial descriptions, reference frame, spatial orientation, map use, spatial judgments, perspective taking

1 Introduction

People can acquire spatial information about an environment through different sources, including navigation, maps, and verbal descriptions. Navigation provides direct experience with the environment. Acquisition from maps or through descriptions is secondary, since others must provide this information [3]. How might mental representations differ when derived from these two different secondary sources? Research has addressed differences between direct experience through navigation and learning through maps, but little research has explicitly examined differences between the secondary sources. All of these media can present the same basic spatial information, but they differ in ways potentially significant for how the information is ultimately represented.

1.1 Spatial Perspective

The spatial perspective that one acquires during learning could influence how one mentally represents the space. Perspective most commonly refers to the viewpoint one takes on the environment, either within the environment (route perspective) or above it (survey perspective). Navigation provides a route perspective and maps give a survey perspective. Spatial descriptions can present one or both perspectives. Research that has examined the relation between acquired representations and perspectives has yielded mixed results. Some studies support perspective-based representation differences [4-8]; other studies have shown no perspective-based differences or have found changes in representation over time [8-10].

The lack of consensus in this research may stem from how perspective is defined. Spatial perspective is not defined solely by viewpoint. A complete definition must also consider the reference system for locating new landmarks, whether the orientation is stable or dynamic, and the amount of information available at a given time. Survey perspectives locate new landmarks with respect to known ones, keep a stable orientation to the environment, and have a substantial amount of information available. Route perspectives use the observer's current location to locate new landmarks, continually change orientation with respect to the environment, and have limited information available at a given time. Because research on spatial perspective generally relies on only a partial list of these features, it is not surprising that conclusions might differ.

Verbal descriptions of space provide a special case of presenting spatial perspective. Verbal descriptions can present either a survey or a route perspective [1, 11]. A route perspective description has all the characteristics of a route perspective, including hypothetical movement through the environment. A survey perspective description has all the characteristics of a survey perspective, save one: because spatial information is related through language, only limited information is available at a given time. This characteristic of survey descriptions is similar to the route perspective.

Unlike navigation and maps, spatial descriptions are not limited to a single perspective. Taylor and Piven [12] found that participants produced route, survey, and mixed perspective descriptions both after studying a map and after learning an environment through navigation. Mixed perspective descriptions took one of two formats, either redundant or alternating. Interestingly, the switches between perspectives were rarely signaled.

A third perspective can also be presented through descriptions, referred to as the gaze perspective. This perspective has elements of both survey and route perspectives. At first glance, the gaze perspective appears similar to the route perspective, because both imply a tour of the layout. However, the route perspective is an implied physical tour; the gaze perspective tours using only the eyes. Gaze tours
maintain a fixed position with respect to the environment and relate object locations to other object locations, like survey descriptions. At the same time, they use a reference frame defined by the speaker, like a route perspective. Ehrlich and Koster [13] found that participants used gaze tours when describing a dollhouse room viewed from the outside. Shanon [14] found similar results for dorm room descriptions; Levelt [15] found the same in descriptions of node-link networks.

1.2 Reference Frames

Presenting spatial information through descriptions creates the possibility of language-based ambiguity. One source of ambiguity is the availability of more than one reference frame. Even a simple conversation between two individuals can lead to reference frame ambiguity. For example, Phil and Kathy stand face-to-face in conversation. During the conversation, Kathy points at, “You have something stuck to the side of your face.” Phil responds, “Which side?” Kathy responds, “the right side.” Fifty percent of the time, Phil probably reaches for the wrong side of his face. Why? The phrase, “the right side,” could be using Kathy’s reference frame or could be using Phil’s. Because they face one another, right using Kathy’s reference frame equals left using Phil’s. Since neither conversant stated the reference frame they were using, there is a fifty-fifty chance of using the correct one.

Ambiguity is most likely to arise when reference frames use the same spatial terms and the frames are misaligned, such as in the example above. When frames are misaligned, does one frame receive priority processing? A pure dominant reference frame seems unlikely. Evidence indicates that reference frame use may be situation dependent. For example, object features, such as movement, make the object’s intrinsic properties more salient [16]. A functional relationship between two objects also influences frame selection [17], as does physical proximity [18]. Reference frame selection also differs depending on whether one is talking to a live conversational partner or is alone [19].

Despite the importance of reference frames, they do not account for all spatial term ambiguity. Interpretation of terms using a single reference frame can also be fraught with ambiguity. Hill [2] discusses reference fields used to impose a relative reference frame on objects without intrinsic sides or asymmetries. The phrase “the keys are to the right of the ball” can be interpreted using the speaker’s reference frame. Hill [2] argues that when interpreting phrases such as this, the speaker sets up an orienting field, generally parallel with his/her own body. However, he points out that different spatial terms may use different orienting fields. The terms left/right generally use an aligned field, such that the keys are to the right side of the ball as defined by the speaker’s right side. The majority of English speakers, however, use a facing field for the terms front/back. For the phrase “the keys are in front of the ball”, most speakers of English would interpret the keys as being between themselves and the ball. If they had been using an aligned field, the keys would be interpreted as being on the far side of the ball. Hill [2] also argues that not all languages use the same alignment of the orienting field for the same terms. Native Hausa speakers generally use an aligned field for their equivalent of front/back. This changes, however, if one object occludes another. Levinson [20] found that speakers of Tzeltal also differ in their assumptions about reference frame interpretation compared to speakers of Western languages, such as Dutch.

Extended spatial descriptions, compared to single sentence descriptions, bring additional challenges for reference frame interpretation. First, the reference frame may change during the description. Levelt [15] asserted that extended spatial descriptions would adopt a single reference frame. This assertion, however, has not been supported in empirical work [12, 19]. Further, Taylor and Tversky [12] found that reference frame shifts are rarely signaled. Second, recipients of extended spatial descriptions must both determine the reference frame used and integrate newly described spatial locations with those previously described. In other words, they must interpret new information while maintaining information about multiple other spatial relations in memory.

2. Present Research

The present work began as a pilot study to investigate the children’s ability to acquire and form spatial representations from different sources. Hence the stimuli take a simple form. The results of pilot testing of university students’ learning an environment from either a description or a map prompted separate research on ambiguity in spatial descriptions, focusing on reference frames and signals for orientation within a reference frame. We report here the work with adults.

The present research had two goals. The first was to directly compare spatial representations derived from different secondary sources. The second was to identify variables that affect the interpretation of and representations acquired from spatial descriptions. Participants learned a description of a six-room house either from a gaze tour description or from a map. We selected a gaze tour description rather than a survey description based on the initial developmental goal of the study. We felt that young children would have difficulty both learning the description and interpreting the canonical locative terms (north, south, east, west). Also, the gaze tour description, because it incorporates the terms right, left, in front, and behind, creates the possibility of linguistic ambiguity, which was a focus of our work.

We also created tasks that required different perspectives. After learning the environment, participants completed two tasks in counterbalanced order. In one, participants imagined being in one of the rooms, facing a particular direction, and pointed to other, out-of-sight rooms. This task took a route, within-environment viewpoint. In the second task, participants created a model of the house using cutouts. This task took a survey, above-environment viewpoint. Both tasks required consideration of multiple relations among multiple locations.

2.1 Experiment 1

In Experiment 1, we compared the performance of participants who learned a spatial layout either from a map or from a verbal description.
Participants. The participants were 48 adults, with equal numbers of males and females. Most were university students who received course credit for their Introductory Psychology course; a few students who were not enrolled in the class also participated and were paid for their time.

Apparatus and Materials. Within the lab space, we constructed a 6 x 6 x 7 foot room using PVC pipes with blue curtains as walls. We used six stuffed animals — pig, cat, rabbit, dog, bear, and frog — to indicate the room participants should imagine themselves to be in on a given trial. During the learning phase, participants sat at a small table. Some of the participants learned the layout of animal’s rooms from a simple map, which consisted of six squares on cardboard with photographs of the animals arranged in the appropriate positions. Square cards with animal photographs attached were also used to reconstruct the layout. The verbal description provided a gaze tour with a viewpoint from above (as in a survey perspective), but using terms right, left, and in front.

Procedures. Participants were assigned to one of two conditions. The Map group learned the room locations from the map. The Description group learned the room locations from a verbal description. The experimenter began by telling the participants that they would learn where six animals lived in six separate but connected rooms. The experimenter also told the participants that they would not be able to see all six rooms; instead, they would see the one room in the middle of the testing space and would have to imagine themselves in a specific animal’s room, with the other rooms around them. The experimenter then showed the participant the stuffed animals, and asked him or her to identify each by name.

For the map condition, the experimenter then produced the cardboard map. One by one, the experimenter placed the animal’s photograph in the middle of each square, saying, “This is the ---’s room. Can you point to where the --- lives?” For each new animal added, the entire sequence was repeated until all six animal pictures were correctly positioned. The experimenter allowed the participant to study the map for as long as he/she wished before removing all study materials from sight.

For the verbal description condition, participants first identified the animals. Next, the experimenter read a description and asked the participant to repeat it from memory. Follow-up questions confirmed whether participants learned the description. For example, if the experimenter said, “The cat lives on the right side of the pig,” the follow-up question was, “Where does the cat live?” This procedure was repeated until the participant could say the entire description twice in succession without error. Across all conditions, the rooms of the house were described/depicted in one of two counterbalanced orders, as indicated below (also see Figure 1).

Description, Order 1
- The cat’s room is on the left side of the rabbit’s room
- The pig’s room is on the left side of the cat’s room
- The dog’s room is in front of the pig’s room
- The bear’s room is on the right side of the dog’s room
- The frog’s room is on the right side of the bear’s room

Description, Order 2
- The cat’s room is on the right side of the pig’s room
- The rabbit’s room is on the right side of the cat’s room
- The frog’s room is in front of the rabbit’s room
- The bear’s room is on the left side of the frog’s room
- The dog’s room is on the left side of the bear’s room

Fig. 1. Layout of the six rooms.

Once the participants expressed confidence in knowing the room locations, we assessed their knowledge using two tasks, presented in counterbalanced order. For the pointing task, participants were instructed to imagine that each time they walked into the constructed room, a different animal lived there and the other animal’s rooms were all around. The experimenter explained that the participant would be asked to point to where other animals lived. Participants then entered the constructed room with the experimenter. For example, they were asked to imagine that they were in the pig’s room. To designate each room, an animal (e.g., the pig) was placed in the middle of the testing room. The participant was asked to stand on one side of the animal, and the experimenter stood on the other side. Thus the participant, the animal, and the experimenter stood side-by-side, facing the same direction. To ensure consistency, all participants faced south, although they were not asked about cardinal directions.

Participants completed 18 points, three from each room. Table 1 shows the set of requested points from each imagined room. For example, when they imagined that they were in the rabbit’s room, the participants were asked to point to the cat, dog, and bear (See also Figure 1). A participant’s point was recorded as falling within one of eight categorical directions—toward one of the four walls or one of the four diagonals (i.e., corners).
Table 1. Direction of points from each imagined room

<table>
<thead>
<tr>
<th>Imagined Room</th>
<th>Requested Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>rabbit</td>
<td>cat</td>
</tr>
<tr>
<td>cat</td>
<td>rabbit</td>
</tr>
<tr>
<td>pig</td>
<td>cat</td>
</tr>
<tr>
<td>dog</td>
<td>rabbit</td>
</tr>
<tr>
<td>frog</td>
<td>cat</td>
</tr>
<tr>
<td>bear</td>
<td>rabbit</td>
</tr>
</tbody>
</table>

The construction task took place at the same table where learning took place. Participants arranged the cutouts (with the animal photographs affixed) to construct the house layout. Participants received the cutouts in a randomized stack. The experimenter recorded the completed configuration.

Results and Discussion. We first assessed whether the participants' performance was affected by how they learned the layout—from the map or from the description. We examined performance on both the pointing and construction tasks.

The primary dependent variable of the pointing task was the total number of correct points, out of 18 possible. The design involved a 2 (Condition: Map or Description) by 2 (Order: Construction Task First Vs. Pointing Task First) by 2 (Sex) ANOVA. The main effect of Condition was significant, $F(1, 40) = 6.44$, $p < .05$. Participants who learned the layout from the map ($M = 11.75$, $SD = 7.42$) pointed more accurately than those who learned from the description ($M = 7.17$, $SD = 5.58$). The interaction between Condition and Sex was also significant. Women ($M = 15.33$, $SD = 5.21$) who in the map condition performed much better than women in the description condition ($M = 6.83$, $SD = 5.76$). The performance of men did not differ based on condition ($M = 8.17$, $SD = 7.74$ for the map condition; $M = 7.5$, $SD = 5.62$ for the description condition). No other main effects or interactions reached significance.

Analyses of participants' constructions measured how well participants preserved spatial relations. We compared where participants placed each photograph in the construction to the correct location. Eighteen participants in the map condition reconstructed the entire configuration correctly, but only 6 participants in the description condition performed this well, $\chi^2(1, N = 48) = 12.00$, $p < .001$. A more detailed examination of construction performance revealed distinct error patterns between conditions. Eight of the description participants, but none of the map participants, reversed the rows, $\chi^2(1, N = 48) = 13.71$, $p < .001$ (see Figure 2); these participants placed the animals from the top row of the correct layout on the bottom row of their reconstructions (and vice-versa). In contrast, the map group errors involved swapping positions of only two animals. Few description group participants made this type of error. The tendency of the description group to reverse the rows is particularly interesting in light of our focus on orientation in the task. In particular, errors that preserved the correct placements of animals within a row but reversed the rows likely stems from misinterpreting "in front."

Fig. 2. Examples of constructions. 2(a) represents a correct construction. 2(b) illustrates a row reversal error. 2(c) illustrates switching of two locations.

The results presented thus far indicate that participants in the description conditions made more errors than those in the map condition. We reasoned that at least some of the description participants' errors might have stemmed from the tendency to interpret "in front" differently than expected [2]. To test this hypothesis, we rescored the pointing task, using the participant's reconstructed map as the correct configuration. In other words, we asked whether participants pointed in a manner consistent with their reconstruction. This adjusted pointing score measured consistency between pointing based on memory and reconstruction of the map based on memory. The analysis revealed a Condition by Sex interaction, $F(1, 40) = 10.41$, $p < .01$. Map-group women ($M = 15.42$, $SD = 5.23$) performed better than description-group women ($M = 9.67$, $SD = 7.48$). In contrast, map group men ($M = 7.92$, $SD = 7.86$) performed worse than description-group men ($M = 14.75$, $SD = 5.17$).

The most important result for the adjusted-total analysis was the lack of a significant main effect for Condition. Although both the original and adjusted pointing data showed a Condition by Sex interaction, a close look at the pointing
scores in each case indicates that men in the description group seemed most prone to misinterpreting the meaning of "in front" in the description. The analysis of the adjusted pointing scores revealed, however, that men in the description group used a consistent mental representation for both tasks.

2.2 Experiment 2

The results of Experiment 1 suggest that participants who learned locations from descriptions were more likely to reverse the rows of the house than were participants who learned from a map. This result supports the conclusion that misinterpretation of spatial terms, in particular the term "in front," contributed to the relatively poor performance of the description group. The concordance between participants' reconstructions and pointing seem to indicate that the description group formed a mental representation of the layout and used this representation in the pointing task. In other words, they performed consistently across tasks.

In Experiment 2 we examined two factors that might influence participants' construal of the term "in front". The first was the position of the animal within the room. In Experiment 1, we placed the animals in a direction that accorded with the ground-level view experienced in the room. In other words, the participant, experimenter, and animal all faced this direction in the room. However, if participants adopted a survey-like perspective, then seeing the animals in this ground-level orientation may have caused confusion. We therefore examined whether participants used the animal's facing direction to interpret orientation information. Accordingly, in Experiment 2, some of the participants saw the animals facing up, as they would be depicted on a map. The remaining participants saw the animals facing in orientation used in Experiment 1. Second, we examined whether providing information about the interpretation of "in front," by using an arrow, would facilitate the formation and use of an accurate mental representation. We used the arrow during the learning phase to show participants our intended interpretation of "in front." Because our focus was on the factors influencing performance in the description condition, all participants in this experiment learned the layout from a description.

Participants. The participants were 72 university students, with equal numbers of men and women. The participants were recruited from the same sources as in Experiment 1. Participants received course credit. In addition, we included a comparison group the 24 participants from the description group in Experiment 1.

Materials. We used a ball, a block, and an arrow constructed out of cardboard to indicate directions. We also used a small bucket to prop the animals face-up in the middle of the blue room. All other materials were identical to those of Experiment 1.

Procedures. All participants learned the configuration from the description used in Experiment 1. The procedures differed from those of Experiment 1 in two ways. First, for half of the participants (the Arrow group), we provided specific interpretations of spatial terms [both right and front] using the arrow, ball, and block. Before reading the description to participants, we placed the block on the right side of the ball and used the arrow to illustrate that "in front" meant toward the participant. Second, we varied the direction that the animals faced when participants were in the room. For half of the participants, we placed the animals facing up, toward the ceiling. The remaining participants saw the animals facing in the direction used in Experiment 1.

Results and Discussion. As in Experiment 1, we first analyzed the pointing task, using the total number of correct points (out of 18) as the dependent variable. In this case, the design of the analyses was a 2 (Arrow: Present or Absent) by 2 (Animal Direction: Same as Participant or Facing Up) by 2 (Order: Pointing Task First or Construction Task First) by 2 (Sex) ANOVA. The main effect of the Animal Direction was significant, F(1, 38) = 15.54, p < .001. Participants who saw the animals facing up (M = 12.58, SD = 5.46) pointed more accurately than those who saw the animals facing in the same direction as themselves (M = 8.15, SD = 5.74). The main effect of Order was also significant, F(1, 38) = 4.07, p < .05. Participants who performed the pointing task first (M = 11.50, SD = 5.56) performed significantly better than those who performed the construction task first (M = 9.23, SD = 6.29). No other main effects or interactions reached significance, including any involving Sex.

Performance on the construction task generally reflected performance on the pointing task. Thirty of the forty-eight participants who saw the animals facing up had correct constructions, but only twenty of those who saw the animals facing in the same direction as themselves had correct constructions. \chi^2(1, N = 96) = 4.18, p < .05. Additional examination of the constructions revealed important error pattern differences between the two groups. For example, participants who saw the animals facing in the same direction as themselves reversed the rows, but only 3 of the participants who saw the animals facing up made this kind of error, \chi^2(1, N = 96) = 4.76, p < .05.

Taken together, these results suggest that many participants who learned from the description had difficulty establishing an orientation within the room and, consequently, interpreting "in front" during learning. This problem is reflected in the relatively large number of row reversals seen during reconstruction. Because seeing the animals facing up helped to reduce these kinds of errors substantially, we conclude that participants may have used the animal's facing direction to orient themselves while pointing.

3. General Discussion

Our results help to shed light on how people interpret and use spatial information acquired from maps and from descriptions. The description group appeared to have had some specific problems that the map group did not; these problems appear to involve, at least in part, reconciling differences in orientation as defined by different perspectives. Participants in this condition learned the environment from one perspective, and then performed tasks that required them to take a different perspective.
influence their interpretation. Further, the presence of different, non-aligned reference frames may influence individuals to consider alternative reference frame interpretations. In other words, there is not a standard case for determining reference frame and spatial orientation selection.

Author Note

This research was funded partly through a Project Varenius Seed Grant from the National Center for Geographic Information Analysis (NCGIA), University of California, Santa Barbara to Holly A. Taylor and David H. Uttal.

References