



ELSEVIER

Cognitive Development 17 (2002) 1157–1183

COGNITIVE
DEVELOPMENT

Structure mapping in spatial reasoning

Merideth Gattis*

Max Planck Institute for Psychological Research, Munich, Germany

Received 1 June 2001; received in revised form 1 April 2002; accepted 1 April 2002

Abstract

Four experiments examined whether spatial reasoning about non-spatial concepts is based on mapping concepts to space according to similarities in relational structure. Six- and 7-year-old children without any prior graphing experience were asked to reason with graph-like diagrams. In Experiments 1 and 2, children were taught to map time and quantity to vertical and horizontal lines, and then were asked to judge the relative value of a second-order variable (rate) or a first-order variable (quantity) represented in a function line. Children's judgments indicated that they mapped concepts to space by aligning similar relational structures: quantity judgments corresponded to line height, and rate judgments corresponded to line slope. These correspondences entail a mapping of first-order concepts to first-order spatial dimensions and second-order concepts to second-order spatial dimensions. Experiments 3 and 4 investigated the role of context in establishing relational structure. Children were taught to map age and rate (Experiment 3) or age and size (Experiment 4) to vertical and horizontal lines, and were then asked to judge the rate or the size represented by a function line. In this context, both rate and size were first-order variables, and children's judgments corresponded to line height, also a first-order variable. The results indicate that spatial reasoning involves a structure-sensitive mapping between concepts and space.

© 2002 Elsevier Science Inc. All rights reserved.

Keywords: Spatial reasoning; Mapping; Graphs; Relational structure

Spatial reasoning is the internal or external use of spatial representations, such as arrays, graphs, and diagrams, to reason. Spatial reasoning may involve reasoning about space itself, as for instance when we compare the lengths of two routes on

* Present address: School of Psychology, Cardiff University, P.O. Box 901, Cardiff, CF10 3YG, UK. Tel.: +44-29-2087-4007; fax: +44-29-2087-4858.

E-mail address: m.gattis@sheffield.ac.uk (M. Gattis).

a map, but it often involves reasoning about non-spatial domains, as for example when we use a graph to reason about time, quantity, or rate. When we use graphs, charts and diagrams to reason about non-spatial concepts, we create mappings between spatial and conceptual information, and use those mappings to generate inferences about the conceptual relations by visually inspecting the observable spatial relations.

1. Mapping concepts to space

This paper concerns the process by which concepts are mapped onto space. One possibility is that concepts are mapped to space according to arbitrary conventions, which must be culturally transmitted, as in the case of written language. For the most part, the rules of transcription for written languages are arbitrary. The written marks representing the sound “nee,” for instance, differ according to language. This arbitrary relationship between phonemes and their visual representation is what makes learning to read such a difficult and lengthy process.

A second possibility is that at least some of the mappings of concepts to space are non-arbitrary and cognitively constrained. One reason to think that this latter possibility may be true is that several studies reported that certain mappings of non-spatial concepts to space are remarkably consistent for both children and adults. [Tversky, Kugelmass, and Winter \(1991\)](#) asked 5- to 11-year-old children from English, Hebrew and Arabic language cultures to place stickers on a square piece of paper to represent three levels of space, time, quantity, and preference. Children from all three language cultures avoided representing increases for any concept in a downward direction. When children represented increases vertically, they typically placed the lowest level of a concept at the bottom of the page, and the highest level of a concept at the top of the page, so that increases moved in an upward direction. This mapping preference, Tversky et al. pointed out, corresponds to a metaphorical association between “up” and “more,” and “down” and “less”: Tversky et al.’s results fit nicely with [Handel, De Soto, and London’s \(1968\)](#) report that adults asked to assign word pairs to spatial locations on a cross-like diagram preferred to assign quantity to the vertical, placing “more” at the upper vertical end and “less” at the lower vertical end.

A second consistency in the mapping of concepts to space is the report of [Gattis and Holyoak \(1996\)](#) that adults assigned “faster” to “steeper” in a graphical reasoning task. In several experiments, adults judged the relative rate of two continuous linear variables using simple line graphs. Graph-like diagrams were constructed which varied the assignments of variables to axes, the perceived cause–effect relation between the variables, and the causal status of the variable being queried. Across all of these conditions, a single factor seemed to account for reasoning performance. People were more accurate at rate judgments when the variable being queried was assigned to the vertical axis, so that a mapping existed between the positive slope of the function line and the positive rate of change. In contrast,

other graphing conventions, such as assigning the causal variable to the horizontal axis, and diagrammatic conventions, such as maintaining the verticality of altitude by assigning it to the vertical axis, had little effect on rate judgments. Gattis and Holyoak concluded that graphical reasoning about rate relations was facilitated by a mapping of slope and rate (“steeper equals faster”), even when it violated other natural mappings (such as a vertical line representing the verticality of altitude).

Results such as those discussed above suggest that some mappings of concepts to space are not arbitrary, because some mappings are more likely to be created than others, and some mappings are more easily learned than others. If the assignments of “more” to “up” and “faster” to “steeper,” are not arbitrary, where do they originate? The next section identifies two possible principles governing the mapping of concepts to space: meaningful associations, and structural similarity.

1.1. Meaningful associations

Consistent mappings may derive from associations, such as the association between “more” and “up.” In the world of experience, increasing the quantity of any solid, such as books, candy, or sand, usually increases its vertical extent as well, leading to a strong association between quantity and vertical extent. The findings of Tversky et al. indicate that such associations indeed influence children’s graphic constructions.

Two factors indicate, however, that association-based mappings are inadequate for explaining mapping patterns in spatial reasoning. When Gattis and Holyoak (1996) contrasted two natural mappings, the iconic mapping of “up” on a vertical line and “up” in the atmosphere against the metaphoric mapping between steeper slope and faster rate of change, the latter mapping exerted a stronger influence on reasoning performance. This finding is significant because it demonstrates that association-based mappings may come into conflict, and when multiple mappings conflict, some mappings reliably take precedence over others. In other words, even if mappings are derived from associations, some kind of coherent system appears to be guiding which mapping is used. Furthermore, the direction and strength of some mapping patterns are not easily explained by experience, such as the association between “steeper” and “faster”: our experience in the physical world is equally likely to lead to association between “steeper” and “slower.” Although walking, riding or driving downhill may lead to an association between “steeper” and “faster,” since steeper hills lead to faster rates of travel downhill, the reverse is equally true for uphill travel, when steeper hills lead to slower rates of travel.

1.2. Mapping relational structure

A second explanation for mapping consistencies is that mappings between concepts and space are based on general constraints governing the mapping process, rather than or in addition to specific associations. An example of such a general constraint is the tendency observed in analogical mapping to map two concepts based

on structural similarities (Gentner, 1983). When asked to compare two problems or concepts, adults tend to map elements of one concept or problem to elements of the other, to map the relations between elements in one concept the relations between elements in the other concept, and to map relations between relations in one concept to relations between relations in the other concept. This sensitivity to relational structure emerges in early childhood. Children are sensitive to the relational structures of perceptual analogy tasks by the age of 4, and will align the relational structures to choose “matches” to complete an analogy by the age of 6 (Kotovsky & Gentner, 1996).

If the tendency to map corresponding relational structures extends beyond semantic concepts and analogical reasoning to include spatial reasoning as well, it could explain children’s tendency to align height and quantity (as reported by Tversky et al., 1991), which are both first-order variables because they are relations between elements, and adults’ tendency to align slope and rate (as reported by Gattis & Holyoak, 1996), which are both second-order variables because they are relations between relations. The following experiments investigated the hypothesis that sensitivity to relational structure influences spatial reasoning by examining how young children with no experience with graphs map first-order and second-order conceptual relations to spatial relations when reasoning with graph-like diagrams. Examining how young children with no graphing experience use spatial diagrams to reason about non-spatial concepts allows us to test whether mapping consistencies are an effect of learning graphing conventions, or of general cognitive constraints on the mapping process. The relational structure hypothesis predicts that even children with no graphing experience should exhibit a tendency to map first-order conceptual relations to first-order spatial relations, and second-order conceptual relations to second-order spatial relations.

In the four experiments presented here, children were given brief introductory training tasks which introduced them to the elements of line graphs, followed by a judgment task which provided the dependent variable. Reasoning about first-order and second-order variables was investigated across different experiments: Experiment 1 examined how children use spatial representations to reason about a second-order variable (rate), Experiment 2 examined how children reason about a first-order variable (quantity), and Experiments 3 and 4 again looked at first-order variables (rate and size), this time investigating the role of context in establishing whether a variable is treated as first- or second-order.

2. Experiment 1

Six- and 7-year-old children without any graphing experience were taught to reason with very simple graph-like diagrams. There were three introductory training tasks whose purpose was to teach children the elements of graphs, followed by a final judgment task which provided the dependent variable for the experiments. In the first and second training tasks children learned to map discrete values of a

dimension across first horizontal and then vertical lines using a procedure based on that of Tversky et al. (1991). In the third training task the children learned to coordinate values from the horizontal and vertical lines “to make a story,” with a procedure similar to that used by Bryant and Somerville (1986). The children in effect learned to map a function, but were taught to think of that function as a story represented by a line. In the final phase of the experiment, the children were asked to judge the rate of an event represented by a particular line.

2.1. Method

2.1.1. Participants

Eighty-four first graders (43 girls and 41 boys) from two elementary schools in Munich, Germany participated in Experiment 1. Children were 6–8 years old (mean age: 6–10; range: 6–0 to 8–0; S.D.: 6 months). Ages for children in each of the four experimental groups were as follows: Time-Up (mean age: 6–10; range: 6–0 to 7–11; S.D.: 5 months), Quantity-Up (mean age: 6–10; range: 6–1 to 8–0; S.D.: 5 months), Time-Down (mean age: 6–10; range: 6–0 to 8–0; S.D.: 6 months), Quantity-Down (mean age: 6–11; range: 6–2 to 8–0; S.D.: 7 months). All children had no prior exposure to graphs. Children’s exposure to graphs, or any diagram or model representing non-spatial concepts spatially, was carefully screened at the level of the teacher, the parents, and each individual child. All who had experience with graphs, timelines, or related models or diagrams were excluded from these experiments.

2.1.2. Design

The experiment consisted of three introductory tasks followed by a final judgment task. The three training tasks were: mapping values to a horizontal line, mapping values to a vertical line, integrating values across these two lines.

There were four experimental groups. Children in the Time-Up group were taught to map increasing quantities to the horizontal line from left to right and increasing times to the vertical line from bottom to top. Those in the Quantity-Up group were taught to map increasing times to the horizontal line from left to right and increasing quantities to the vertical line from bottom to top. Those in the Time-Down group were taught to map increasing quantities to the horizontal line from left to right and increasing times to the vertical line from top to bottom. Those in the Quantity-Down group were taught to map increasing times to the horizontal line from left to right and increasing quantities to the vertical line from top to bottom. These four groups formed a 2×2 design, with the two variables being: (1) mapping time on the horizontal and quantity on the vertical versus vice versa and (2) mapping increases on the vertical line upwards versus downwards.

2.1.3. Procedure

Each child was tested individually. The experimenter first asked some simple questions to create a more familiar environment, and then explained that she would

show some lines on pieces of paper, tell some stories with those lines, and ask a few questions. Complete versions of all stories (translated from German) can be found in [Appendix A](#).

2.1.3.1. Part 1. Preliminary training tasks. The experiment began with three brief training tasks to introduce children to the elements of graphs: mapping values to a horizontal line, mapping values to a vertical line, and integrating values across these two lines. In the first task, each child was shown a piece of paper with a L-shaped frame (similar to a Cartesian graph, see [Fig. 1](#)). The experimenter taught the child to map three discrete values of a variable (time for half of the children and quantity for the other half) to the horizontal line by placing three stickers on the line. This task was repeated three times with different stories. For the first story, the experimenter placed the first two stickers (i.e., stickers for breakfast time and lunch time), and asked the child to place the third sticker (i.e., a sticker for dinner time). For the second story, the experimenter placed the first sticker, and asked the child to place the remaining two stickers. For the third story, the child was asked to place all three stickers. The three temporal stories were about meals in a day, activities in a day, or steps in a familiar procedure, tooth brushing. The three quantity stories were about different amounts of books, candy, and sand.

The second training task was mapping values of a variable to the vertical line. This task involved the same sticker-modeling procedure as the first task. Children who had learned to map time to the horizontal line during the first task were taught

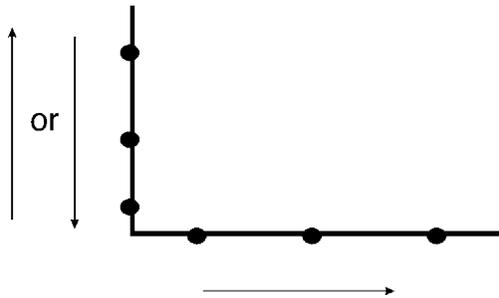


Fig. 1. Simple L-shaped frames were used in the first two training tasks for all four experiments. In the first training task children were taught to map increasing values of a variable to the horizontal line by placing stickers from left to right. In the second training task children were taught to map increasing values of a variable to the vertical line by placing stickers, with half of the children taught to map increases in an upward direction and half of the children taught to map increases in a downward direction. The arrows above represent direction of increase. Assignment of variables to axes was also manipulated experimentally, so that in each experiment, half of the children were taught to map one variable to the horizontal line and a different variable to the vertical line, and half of the children were taught the opposite assignment. The two variables assigned to the axes were time and quantity in Experiments 1 and 2, age and rate in Experiment 3, and age and size in Experiment 4.

to map quantity to the vertical line, and those who had learned to map quantity to the horizontal line were now taught to map time to the vertical line. Half of the children from each group were taught to map increases in an upward direction (placing the smallest value or earliest event at the bottom of the line) and half were taught to map increases in a downward direction (placing the smallest value or earliest event at the top of the line). This manipulation was designed to allow a test of whether children's responses in the final judgment task corresponded to the height or the slope of a line (see below).

In the third training task, the experimenter taught children to coordinate values from the horizontal and vertical lines and to place a sticker at those points representing an integrated time–quantity value. The experimenter first demonstrated how to find the intersection points by reminding the child of the sticker placements from the previous tasks, placing two small birds on each of the points representing the smallest value on each line, and showing the child how to find where the two birds will meet if they each fly in a straight line. This task involved three intersections, and after finding all three intersections, the child was shown that the intersections also form a line. The entire intersection task was then repeated three times.

Finally, the experimenter showed the child a new diagram with a line already drawn to represent the intersection of time and quantity. If the child had been asked to map increases in an upward direction, the line was drawn 60° from the vertical, as in Fig. 2A. If a child had been asked to map increases in a downward direction, the line was drawn 110° from the vertical, as in Fig. 2B. The experimenter then told two stories about a continuous event represented by this line (a dump truck dumping sand and a bathtub filling with water — see Appendix A for stories), and taught the child to place three stickers representing the intersection of two values each. During the first story, the experimenter modeled the first intersection points and asked the child to place the last two, and during the second story the experimenter asked the child to place all three stickers.

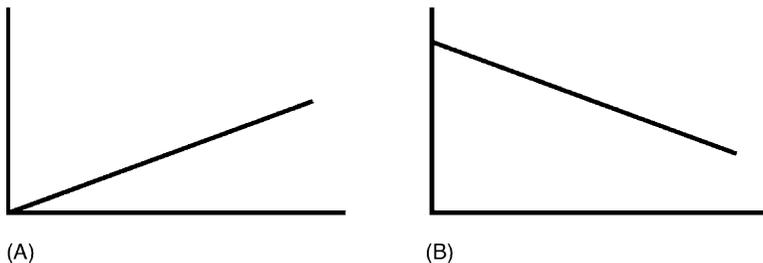


Fig. 2. Experiments 1 and 2 had a third training task, in which children were shown how the values from each variable could be integrated to form a function line. Children in the Time-Up and Quantity-Up groups saw a function line that began in the lower left corner of the frame (A), while children in the Time-Down and Quantity-Down groups saw a function line that began in the upper left corner (B).

2.1.3.2. Part 2. Judgment task. In the critical judgment task, children were shown a diagram of the L-shaped frame with two function lines. One function line was exactly the same as in the previous diagram, and the other data line was 10° from it (70° in the upward-mapping condition, or 100° in the downward mapping condition). The diagrams given to all four experimental groups are illustrated in Fig. 3. Children in the Time-Up and Quantity-Up groups, received a diagram with two function lines that both began in the lower left corner of the diagram and sloped upwards. Children in the Time-Down and Quantity-Down groups received a diagram with two function lines that both began in the upper left corner and sloped downwards.

The experimenter told participants, “Here we have two story lines. That means that a similar story happened twice. Imagine I was taking a bath on two different days: one day is one line, and the other day is the other line. Each line stands for

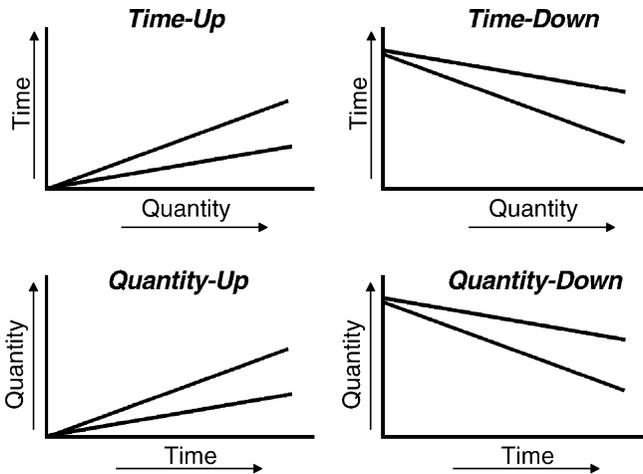


Fig. 3. The four conditions in Experiments 1 and 2. Children taught to map increases along the vertical in an upward direction received upward-sloping diagrams in the judgment phase, as illustrated in the two graphs on the left. Children taught to map increases along the vertical in a downward direction received downward-sloping diagrams in the judgment phase, as illustrated in the two graphs on the right. Note that when increases mapped upward, the upper line was the steeper of the two (seen in the two graphs on the left), but when increases mapped downward, the upper line was the shallower of the two (seen in the two graphs on the right). The illustrations here are also labeled for axes assignment — two groups were taught in the training phase to map quantity to the horizontal and time to the vertical, as seen in the top half of the figure, and two groups were taught the opposite mapping, as seen in the bottom half of the figure. The actual graphs that children received were not labeled, they are labeled here for illustration purposes only. The experimenter pointed at the upper line and asked children to make a judgment about the rate (Experiment 1) or quantity (Experiment 2) represented by that line. The graphs that children received in Experiments 3 and 4 were identical, but children were taught to map age and rate (Experiment 3) or age and size (Experiment 4). In Experiments 3 and 4 the experimenter pointed at either the upper or lower line, and asked the child to judge the rate (Experiment 3) or the size (Experiment 4) represented by that line).

filling up the bathtub on a different day. One day, I turned on the water full power, the tub fills up much faster. Another day, I turned on the water just a little, so that the tub fills up more slowly.” The experimenter pointed at the end of the upper line, and asked children which event was represented by that line, “Look at this line. Does this line stand for the time it happened faster or for the time it happened slower?” The experimenter then produced a new but identical sheet of paper, and repeated the task with a story about two dump trucks dumping sand at different rates.

The vertical direction of increase and the resultant diagrams shown in Fig. 3 were manipulated to test whether children’s rate judgments would correspond to line height or to line slope. The logic was as follows. Because of the children’s lack of exposure to graphs and the rules of graphing, combined with the simple and preliminary nature of the training procedure, children were not expected to answer accurately according to the normal conventions of graphing, but rather to map the probed concept (rate in this experiment) to a relevant spatial relation from the diagram. The graphs were constructed in such a way that there were two salient differences between the lines: height and slope. If children were to choose the height of the line as the relevant spatial relation to which rate should be mapped, children in all groups would be expected to answer “faster” to the probed upper line, because in all four cases, the upper line has by definition greater height. The structure mapping hypothesis predicts instead that children would choose the slope of the line as the relevant spatial variable, because both rate and slope are relations between relations, or second-order variables. If children were to choose the slope of the line as the relevant spatial relation, children in the Time-Up and Quantity-Up conditions would be expected to answer “faster” to the probed upper line, while children in the Time-Down and Quantity-Down conditions would be expected to answer “slower” to the probed upper line — because when the function lines slope up, the upper line has greater slope, but when the function lines slope down, the upper line has lesser slope.

2.2. Results

2.2.1. Preliminary training tasks

Children in all experimental conditions learned to map and integrate values from horizontal and vertical lines quickly, and with virtually no errors. No performance differences were found to indicate preferences for representing particular concepts across particular dimensions of space (for instance, a preference for representing time horizontally).

2.2.2. Rate judgment

Because the main question of this study is whether children exhibit consistent biases in spatial reasoning, and what those biases might be, judgments were scored for consistent patterns, rather than accuracy as defined by normal conventions of graphing. The results are reported as frequencies in Table 1. Answers from children

Table 1
Rate judgments for each condition of Experiment 1

Mapping direction and axis assignment	Number reporting “faster” for both judgments	Number reporting “slower” for both judgments	Number of inconsistent judgments	Total number in group
Vertical increased upward				
Time on vertical	9	5	7	21
Quantity on vertical	15	1	5	21
Vertical increased downward				
Time on vertical	4	12	5	21
Quantity on vertical	2	14	5	21

shown diagrams with upward-sloping function lines (Time-Up and Quantity-Up, shown on the left side in Fig. 3) are shown in the top half of the table, and answers from children shown diagrams with downward-sloping function lines (Time-Down and Quantity-Down, shown on the right side in Fig. 3) are shown in the bottom half of the table. Rate judgments were not determined by whether time or quantity was mapped vertically, $\chi^2(2, N = 84) = .84, P > .05$. For this and all of the following experiments, responses were therefore collapsed across the two horizontal and vertical variable assignments (horizontal quantity and vertical time, and horizontal time and vertical quantity). Rate judgments were determined instead by whether increases mapped upward or downward, $\chi^2(2, N = 84) = 24.45, P < .01$. When increases mapped upward, children were more likely to answer that the probed upper line represented the “faster” event. When increases mapped downward, children were more likely to answer that the probed upper line represented the “slower” event. Thus, children’s rate judgments corresponded to the slope of the line, as predicted by the structure mapping hypothesis.

Rate judgments did not correspond to the height of the line. The experimenter always asked for a judgment of the higher of the two lines, but children’s responses varied with the vertical direction of increase, indicating that slope, rather than height, influenced children’s rate judgments.

3. Experiment 2

Experiment 2 was identical to Experiment 1, with one critical difference. Children were asked to judge the value of quantity, a first-order variable, rather than the value of rate, a second-order variable. This difference was expected to lead to a different pattern of judgments, because it was expected that sensitivities to relational structure would lead children to an appropriately similar first-order variable from the spatial relations in the graph. This variable was expected to be the height of the lines, leading to similar responses from all groups, rather than the diverging responses seen in Experiment 1.

3.1. Method

3.1.1. Participants

Thirty-eight first graders (22 girls and 16 boys) from another elementary school in Munich, Germany participated in Experiment 2. Children were 6–8 years old (mean age: 7–2; range: 6–0 to 8–0; S.D.: 5 months). Ages for children in each of the four experimental groups were as follows: Time-Up (mean age: 7–0; range: 6–0 to 8–0; S.D.: 6 months), Quantity-Up (mean age: 7–1; range: 7–0 to 8–0; S.D.: 4 months), Time-Down (mean age: 7–3; range: 7–0 to 8–0; S.D.: 5 months), Quantity-Down (mean age: 7–4; range: 7–0 to 8–0; S.D.: 6 months). Children's exposure to graphs was screened in the same manner as Experiment 1, and no children had prior experience with graphs.

3.1.2. Design

The 2×2 design was identical to the design of Experiment 1.

3.1.3. Procedure

The procedure of Experiment 2 was identical to that of Experiment 1, with the crucial difference that in the final judgment task, children were asked to judge quantity produced by an event, rather than the rate of an event. For both stories, the experimenter pointed at the end of the upper line, and said, "Look at this line. Does this line stand for the time when there was more or for the time when there was less?"

3.2. Results

No performance differences were found on the training tasks. As in Experiment 1, children in all experimental conditions learned to map and integrate values from horizontal and vertical lines quickly, and with virtually no errors. Children's quantity judgments are presented as frequencies in Table 2, collapsed across the two variable-to-axes assignments (i.e., vertical time and horizontal quantity versus vice versa), but with the upward-increasing diagrams again represented in the top half of the table and the downward-increasing diagrams again represented in the bottom half of the table. The majority of children in all conditions reported that the probed upper line represented a greater quantity ("more"). Because of the low expected frequencies, a Fisher Exact Probability test was used to test whether there was a significant difference between the conditions. There was no difference

Table 2
Quantity judgments for each condition of Experiment 2

Mapping direction	Number reporting "more" for both judgments	Number reporting "less" for both judgments	Number of inconsistent judgments	Total number in group
Increases mapped up	16	1	1	18
Increases mapped down	11	2	7	20

($P = .33$, Fisher Exact Probability test, Siegel, 1956). Thus, whereas the rate judgments in Experiment 1 corresponded to slope, the quantity judgments in Experiment 2 corresponded to height.

4. Experiment 3

The results of Experiments 1 and 2 are consistent with the hypothesis that spatial reasoning involves an alignment of structurally similar relations between concepts and space. In both experiments, children's judgments of the value of a conceptual variable represented in a diagram corresponded in level of complexity to a structurally appropriate spatial cue. In Experiment 1, children's judgments of rate, a second-order variable, corresponded to line slope, also a second-order variable. In Experiment 2, children's judgments of quantity, a first-order variable, corresponded to line height also a first-order variable.

Experiments 3 and 4 further explored the structure mapping hypothesis by examining the role of context in establishing relational structure. Whereas in Experiment 1 rate had been introduced as a second-order relation composed of two further variables, time and quantity, in Experiment 3 rate was introduced as a first-order variable. In a procedure similar to that used in Experiments 1 and 2, 6- and 7-year-old children without any prior graphing experience were taught to construct linear mappings of age and rate. In the judgment phase of the experiment, the children were told a story about two imaginary animals, and were asked to judge the rate of travel of one animal represented by a particular function line. In this context, rate was a first-order variable. If the relational structure of a concept is determined by context, and if that contextually determined relational structure determines the mapping of concepts to space in spatial reasoning, children's judgments would be expected to correspond to line height, also a first-order variable, rather than line slope, a second-order variable.

4.1. Method

4.1.1. Participants

Seventy-nine first graders (30 girls and 49 boys) from Munich elementary schools participated in Experiment 3. All children had no prior exposure to graphs, timelines, or related models or diagrams. Children were 6–8 years old (mean age: 7–0, range: 6–0 to 8–0, S.D.: 4 months). Ages for children in each of the four experimental groups were as follows: Age-Upper (mean age: 6–11; range: 6–0 to 8–0; S.D.: 5 months), Rate-Upper (mean age: 6–11; range: 6–0 to 8–0; S.D.: 5 months), Age-Lower (mean age: 6–11; range: 6–0 to 7–8; S.D.: 5 months), Rate-Lower (mean age: 7–1; range: 7–0 to 8–0; S.D.: 3 months).

4.1.2. Design

The design was similar to the design of Experiments 1 and 2, with one important difference, that during the judgment task children were either asked to judge the

rate represented by the upper line (as in Experiments 1 and 2) or the lower line (which was not probed in Experiments 1 and 2). Whereas in Experiments 1 and 2 the experimenter had probed only the upper function line, in Experiment 3 the experimenter probed the upper line for half of the children, and the lower line for half of the children. The assignment of variables to vertical and horizontal lines was considered a counterbalancing variable: half of the children were taught to map age to the horizontal line and rate to the vertical line, and half were taught the opposite assignment, but these groups were collapsed for the analyses of results. As in Experiments 1 and 2, half of the children were taught to map increases along the vertical line in an upward direction and half were taught to map increases in a downward direction. The result was a 2×2 design, with direction of vertical mapping as the first variable (upward vs. downward) and line probed in judgment task as the second variable (upper or lower line).

4.1.3. Procedure

As in Experiments 1 and 2, the procedure consisted of two parts: preliminary training tasks that served to introduce children to the elements of graphs, and the final judgment task, which provided the dependent variable.

4.1.3.1. Part 1. Preliminary training tasks. The preliminary training tasks were similar to those in Experiments 1 and 2, with the primary difference that the integration tasks used in Experiments 1 and 2 were replaced by a brief explanation at the beginning of the judgment task about how function lines combine the values represented on the horizontal and vertical lines. Children's performances in Experiments 1 and 2 suggested that integration was not as difficult a task as had been supposed, and that such a lengthy integration procedure was not necessary. A further pilot study confirmed that the integration tasks in Experiments 1 and 2 were not responsible for the pattern of results reported in Experiments 1 and 2. In addition, in Experiment 1 the integration procedure served to emphasize to children that rate was a second-order variable because it was a relation between two other relations, a temporal relation and a quantity relation. In Experiment 3 rate was presented to children as a first-order variable, and the component relations that might determine rate (such as for instance, time and distance) were not mentioned. This also made the lengthy integration task less essential.

The first preliminary training task was mapping values to a horizontal line. A child was shown a piece of paper with an L-shaped figure drawn on it (as in Fig. 1). The experimenter directed the child's attention to the horizontal line, and taught the child to map three values of a variable (either age or rate) to this line by placing stickers on the line, as in Experiments 1 and 2. This task was repeated twice with two different stories for the relevant variable. The age stories were about the different ages of people, and the different ages of dogs. The size stories were about the different sizes of houses, and the different sizes of dogs. The rate stories were about the different rates of travel of different vehicles, and the different rates of travel of different dogs (complete versions of all stories are in

Appendix B). The experimenter first modeled the mapping procedure for two values, as in Experiments 1 and 2, and asked the child to map the third value. When the task was repeated with a new story, the experimenter asked the child to map all three values.

The second preliminary training task was mapping values to a vertical line. Using the same sticker-modeling procedure, the children who had mapped rate or size to the horizontal line were now asked to map age to the vertical line, and those who had mapped age to the horizontal line were now asked to map rate or size to the vertical line.

4.1.3.2. Part 2. Integrating values and judgment task. In the final task, children were shown a diagram of the L-shaped frame with two data lines drawn inside it. Children who had been taught to map vertically ascending values saw two function lines that sloped upward from the lower left corner, at 60 and 70°, as seen in the left column of Fig. 3 (with the difference that the concepts assigned to the axes were age and rate rather than time and quantity). Children who had been taught to map vertically descending values saw two function lines that sloped downward from the upward left corner, at 100 and 110°, as seen in the right column of Fig. 3 (again with the difference that the concepts assigned to the axes were age and rate rather than time and quantity).

The experimenter told the child, “Remember that this line (pointing to the vertical or horizontal axis) tells us about age. This is where you placed stickers for the different ages. And this line (pointing to the other axis) tells us about rate. This is where you placed the stickers for different rates. When we combine age and rate we can make stories — and the last two lines left over are our story lines (pointing to the two function lines). These lines tell stories, and there are two of them. That means that a similar story happened twice.” The experimenter then said, “Those two stories are about two animals, and how they get faster as they get older. You’ve never seen these two animals before, and they don’t look like any animal you’ve ever known. One is called a chimera and one is called a xyrus. They both get faster as they get older, but one of them does it more than the other.” The experimenter then pointed at the end of either the upper or lower line, and said, “Look at this line. Does this line stand for the one that’s faster or slower?” In Experiments 1 and 2 each child had been asked to make two judgments, but in Experiment 3 each child made only one judgment.

4.2. Results

As with Experiments 1 and 2, judgments were scored for consistent patterns rather than accuracy as defined by the rules of graphing. The results are reported as frequencies in Table 3. Answers from children who were asked to judge the rate represented by the upper function line are shown in the top half of the table, and answers from children who were asked to judge the rate represented by the lower function line are shown in the bottom half of the table. Rate judgments were not

Table 3
Rate judgments for each condition of Experiment 3

Mapping direction	Number reporting “faster”	Number reporting “slower”	Total number in group
Upper line			
Increases mapped up	16	3	19
Increases mapped down	18	2	20
Lower line			
Increases mapped up	6	14	20
Increases mapped down	6	14	20

determined by whether increases along the vertical mapped upward or downward. Because judgments for the two mapping directions were virtually identical, these were combined and judgments of the upper and lower line were compared. Of the 39 children asked to judge the rate represented by the upper line, 34 responded that the line represented the faster animal. Of the 40 children asked to judge the rate represented by the lower line, 28 responded that the line represented the slower animal. This difference was highly significant, $\chi^2(1, N = 79) = 26.6, P < .01$.

These results contrast with the results of Experiment 1, in which children were also asked to make rate judgments. Whereas in Experiment 1, children’s rate judgments were influenced by the direction of mapping increases along the vertical line, here children’s rate judgments were not influenced by direction of increase, but rather by whether the upper or lower line was probed during the judgment task. In other words, in Experiment 1, children’s rate judgments corresponded to line slope, and in Experiment 3, children’s rate judgments corresponded to line height. The critical difference between Experiments 1 and 3 is that in Experiment 1, rate was presented as a relation composed of two other relations, a temporal relation and a quantitative relation, so that rate was a second-order variable, while in Experiment 3, rate was presented as a relation along a dimension, or in other words a first-order variable. These results are consistent with the proposal that reasoning with spatial representations relies on a structure-sensitive mapping of concepts to space, so that first-order conceptual variables (i.e., relations along a dimension) are mapped to first-order spatial variables, and second-order conceptual variables (i.e., relations between relations) are mapped to second-order spatial variables.

5. Experiment 4

The purpose of Experiment 4 was to allow a comparison between children’s judgments of rate in Experiment 3 and children’s judgments of another first-order variable, size, using the identical procedure. For this reason, Experiment 4 was nearly identical to Experiment 3, with the simple difference that children were taught to map age and size to vertical and horizontal lines, and were then asked to judge the size of an imaginary animal represented by a function line. Size

was a first-order variable, and therefore children's judgments were expected to correspond to line height, also a first-order variable.

5.1. Method

5.1.1. Participants

Eighty-one first graders (47 girls and 34 boys) participated in Experiment 4. All children had no prior exposure to graphs, timelines, or related models or diagrams. Children were 6–8 years old (mean age: 7–1; range: 6–0 to 8–0; S.D.: 4 months). Ages for children in each of the four experimental groups were as follows: Age-Upper (mean age: 7–0; range: 6–0 to 8–0; S.D.: 4 months), Size-Upper (mean age: 7–2; range: 7–0 to 8–0; S.D.: 5 months), Age-Lower (mean age: 7–1; range: 6–0 to 8–0; S.D.: 5 months); Size-Lower (mean age: 7–1; range: 7–0 to 8–0; S.D.: 3 months).

5.1.2. Design

The 2×2 design was identical to that of Experiment 3.

5.1.3. Procedure

The procedure was identical to that of Experiment 3, with the difference that children were given training and judgments about age and size rather than age and rate.

5.1.3.1. Size judgment. In the judgment task, the experimenter asked the children to judge the size of imaginary animals represented in a graph. The experimenter told a story about the animals, then pointed at the end of either the upper or lower line, and said, "Look at this line. Does this line stand for the one that's bigger or smaller?"

5.2. Results

The results are reported as frequencies in [Table 4](#), following the same format as [Table 3](#). Judgments of the upper line are shown in the top half of the table,

Table 4
Quantity judgments for each condition of Experiment 4

Mapping direction	Number reporting "more"	Number reporting "less"	Total number in group
Upper line			
Increases mapped up	18	3	21
Increases mapped down	18	2	20
Lower line			
Increases mapped up	2	18	20
Increases mapped down	2	18	20

and of the lower line are shown in the bottom half of the table. Rate judgments were not determined by whether increases along the vertical mapped upward or downward. Judgments for both mapping directions were virtually identical, and so these numbers were combined and judgments of the upper and lower line were compared. Of the 41 children asked to judge the size of an animal represented by the upper line, 36 responded that the line represented an animal that was “bigger.” Of the 40 children asked to judge the size of an animal represented by the lower line, 36 responded that the line represented an animal that was “smaller.” This difference was highly significant, $\chi^2(1, N = 81) = 49.3, P < .01$.

These results are similar to the results of Experiments 2 and 3. In Experiment 2, children were asked to judge the quantity represented by a line, in Experiment 3, children were asked to judge the rate represented by a line, and in this experiment children were asked to judge the size represented by a line. These three variables, quantity, rate, and size, were all presented as relations along a single dimension, or first-order variables. In all three experiments, children’s judgments corresponded to the height of the line.

6. General discussion

These experiments examined the origins of a remarkable human ability — the capacity to reason with conceptual information presented spatially. Whereas many aspects of reasoning are understood well enough to build computer models of human performance, no current theory of reasoning explains why humans are so good at reasoning spatially. Previous work has tended to focus on why spatial representations facilitate recognition and search (e.g., [Larkin & Simon, 1987](#)), and left unaddressed the question of why spatial representations are so powerful for reasoning and inference. The four experiments presented in this paper investigated an hypothesized constraint on reasoning that could explain why spatial reasoning is fast and flexible, but also yields intelligent inferences. This constraint is the mapping of relational structure between concepts and space.

Six- and 7-year-old children with no formal training in graphing were given a brief orientation about the elements of graphs, and asked to make judgments about the value represented by one of two sloping function lines. The training and graphs were constructed in such a way that for the judgment task, half of the children saw a graph with two upward sloping function lines, and half of the children saw a graph with two downward sloping function lines. In the upward-sloping graph, the higher of the two lines had both greater height and greater slope, while in the downward-sloping graph, the higher of the two lines had greater height but lesser slope. The graphs were constructed in this way to test whether children’s judgments corresponded to one of two perceptual cues, height and slope, or whether the judgments were random. The type of judgment made was varied across experiments, but the same diagrams were used in all four experiments.

In the first two experiments, children's judgments about rate were compared with judgments about quantity. It was predicted that sensitivity to relational structure would lead children to map a relation along a single dimension, such as quantity, to a relation along a single spatial dimension, such as height, and to map a relation between relations, such as rate, to a spatial relation between relations, such as slope. The results were consistent with this prediction.

In Experiment 1, children's rate judgments were influenced by the direction of mapping increases along the vertical line. Direction of increase appeared to influence rate judgment because of the relationship between direction of increase and the relative slope of the two data lines. When increases mapped upward, the upper line was steeper than the lower line, and when the experimenter probed the upper line, children were more likely to report that it represented a faster event. When increases mapped downward, the upper line was the shallower of the two, and children's answer patterns were reversed: when the experimenter pointed at the upper line, they reported that it represented a slower event.

In Experiment 2, children's quantity judgments were not influenced by the direction of mapping increases along the vertical line. Children in both mapping conditions reported that the probed upper line represented "more." In other words, children's responses were consistent with a mapping of quantity to height, rather than slope.

Experiments 3 and 4 examined how context influences relational structure and thereby the mapping between concepts and space. In Experiment 3, children were asked to make judgments about rate and in Experiment 4 children were asked to make judgments about size, but both rate and size were introduced as first-order variables. In this context, both rate and size were predicted to be mapped to height, a first-order variable, rather than slope, a second-order variable. This was the case. Rate and size judgments corresponded to the height, but not the slope of a line.

The important difference between Experiments 1 and 3 is the conceptual context in which rate was introduced. In Experiment 1, rate was introduced as a second-order relation composed of changes in quantity and changes in time. In contrast, in Experiment 3, rate was introduced as a first-order relation, and no component variables such as time and quantity or time and distance were mentioned. Strikingly, in both experiments, children's judgments of rate were consistent with the corresponding level of spatial structure. In Experiment 1, children's rate judgments were consistent with the slope of the probed line (a second-order relation), and in Experiment 3, children's rate judgments were consistent with the height of the probed line height (a first-order relation). Together the judgment patterns found in Experiments 1 and 3 indicate a remarkable sensitivity to structural similarity in spatial reasoning.

This structure-driven mapping in spatial reasoning may reflect some of the same constraints influencing structure-mapping in analogical reasoning (Gentner, 1983, 1988). While analogical reasoning is characterized by the mapping of similar conceptual structures, spatial reasoning is characterized by mapping conceptual structures to spatial structures, also based on structural similarity. Spatial reasoning

is, in this sense, much like reasoning by analogy: new knowledge is inferred by mapping an abstract concept, the target, onto a familiar and concrete source domain, namely space — an ideal source domain because it is well-learned, and highly constrained due to its limited dimensionality.

Spatial reasoning differs from analogical reasoning, however, in the type of structures being mapped, and therefore in the level at which similarity can be defined: because space is perceptual but many concepts are not, defining similarity between concepts and space often seems more abstract than defining similarity between two concepts. The results reported here demonstrate that at least one type of similarity influencing spatial reasoning is relational structure: in these experiments, first-order relations such as quantity and height (in Experiment 2), rate and height (in Experiment 3), and size and height (in Experiment 4) were mapped together, and second-order relations such as rate and slope (in Experiment 1) were mapped together. Young children thus appear to distinguish relations between elements from relations between relations, and to map concepts to spatial representations accordingly. This finding accords with [Tversky's \(1995\)](#) observation that many graphic depictions involve a mapping of elements to elements and relations to relations, and that young children asked to create notational systems observe a similar rule. Recent studies of diagrammatic reasoning by adults provide further evidence that interpretations of novel diagrams rely on a mapping of elements to elements and relations to relations ([Gattis, 2001, 2002](#)).

6.1. Spatial reasoning and children's understanding of functional relationships

An important assumption of the structure-mapping interpretation of the current results is that 6- and 7-year-old children are capable of reasoning about functional relationships. Support for the view that young children are capable of reasoning about functional relationships is found in a variety of studies with young children. Piaget and his colleagues concluded from several studies that preschool children have an intuitive “logic of functions” which is qualitative in nature and is the basis of mature functional reasoning ([Piaget, Grize, Szeminska, & Bang, 1968/1977](#)). More recent developmental studies also indicate that by the age of five, children have a basic understanding of time, distance, and speed, and are able to integrate two of those dimensions to reason about a third ([Halford, 1993; Wilkening, 1981](#)).

6.2. Spatial reasoning and graphing conventions

The results reported here suggest that spatial reasoning is influenced by general constraints in reasoning, and that these constraints precede the learning of graphing conventions. The children in these experiments had no formal instruction in graphing, and but when asked to interpret the meaning of function lines in graph-like diagrams, answered in highly consistent ways. These consistencies were not necessarily correct according to all the rules of graphing, but they do reflect a basic principle of graphing, which is that conceptual dimensions are mapped to spatial

dimensions according to relational structure. This simple principle of graphing is represented in such complex rules as the rate of change of the dependent variable equals the change in y over the change in x , but it can also be expressed in simple mappings, such as “steeper equals faster” and “shallower equals slower.” The consistency observed here that rate (as a second-order variable) is mapped to slope is similar to adults’ judgment patterns in a more complex and realistic graph interpretation task. Adults asked to make rate judgments about line graphs also make the mappings, “steeper equals faster” and “shallower equals slower,” even when the graph is constructed in such a way that this simple mapping leads to an incorrect answer (see [Gattis & Holyoak, 1996](#)).

6.3. *Spatial reasoning in educational contexts*

In contrast to the facility for spatial reasoning demonstrated by young children in these experiments without any prior graphing instruction, educational researchers have documented numerous failures in graphing performances of school children subsequent to extensive graphing instruction ([Leinhardt, Zaslavsky, & Stein, 1990](#); [McDermott, Rosenquist, & VanZee, 1987](#)). [Leinhardt et al. \(1990\)](#) review many studies of school children’s performance in graph interpretation and construction, in which they note a myriad of common errors in prediction, classification, translation and scaling tasks. These two contrasting pictures of children’s graphing abilities pose an apparent paradox. If spatial reasoning, including reasoning with graphs, is governed by fundamental cognitive constraints, why do young children and even college students often perform poorly on tests of graphing skills? As [Leinhardt et al.](#) point out, most tasks studied in educational settings involve either formal knowledge (i.e., the definition of a function), specific experience (i.e., translation between the algebraic notational system and the Cartesian coordinate system), or construction, which is a more difficult cognitive process than interpretation ([Bates, 1993](#); [Savage-Rumbaugh, 1993](#)). In contrast, the task used in these experiments was a qualitative interpretation task. [Leinhardt et al.](#) note that qualitative interpretation tasks are rare in the mathematics curriculum, but are easier than most graphing tasks, in part because they rely on processing of global features. Some researchers argue that qualitative interpretive tasks ought to be the initial step in graphing instruction ([Bell & Janvier, 1981](#)), and the results reported here indicate that qualitative interpretation tasks may indeed be an excellent introduction to graphing for young children.

Acknowledgments

This research was supported by the Max Planck Society. I thank Tylor Hagerman, Marija Kulis, Denise Parks, and Felicitas Wiedermann for their assistance in preparing and running these experiments, the teachers and children of the Grundschulen Bad-Soden-Strasse, Bayernplatz, Farinelli-Strasse, Klenze-Strasse, Simmern-Strasse, and Torquato-Tasso-Strasse in Munich for their

participation, and Peter Bryant, Dedre Gentner, Keith Holyoak, Barbara Tversky, and Michael Waldmann for valuable comments on this paper.

Appendix A. Cover stories for Experiments 1 and 2

A.1. Time stories

A.1.1. Meals

I want you to think about time. For example, think about the meals in a day. In the morning, you get up and you have breakfast, for example cornflakes or Müsli or your Nutella bread. After you come home from school, you usually have lunch, yummy things like Pizza or Spaghetti. And in the evening, after doing your homework and playing, you have dinner. I've got some stickers here. I'm going to put this sticker here on the line to represent breakfast time (experimenter places sticker near one end of the line). So, this is breakfast, okay? And the day goes on and on and on, time is passing by, and then, it's lunchtime. This is the sticker for lunchtime and I put it right here (experimenter places sticker near the middle of the line). Now it's your turn: Can you put the sticker on the line for dinnertime?

A.1.2. Activities

There are also different things you do during the day: for example, you have to get up and out of your warm bed in the morning when your mum wakes you up. Later in the day, after lunch, you have to do your homework, and at night after dinner, you have to go to bed to sleep. This is our sticker for getting up in the morning, and I'm going to put it right here on the line (experimenter places sticker near one end of the line, in the same place as in the previous task). Now, can you put the sticker for doing your homework in the afternoon? And now, I want you to put the sticker for going to bed at night on that line.

A.1.3. Tooth brushing

You know, there are certain things in your life you have to do always in the same order. No matter whether you are 6 or 60 years old, you always do that in a special order. For example, brushing your teeth. First, you have to take your toothbrush, then, you put the toothpaste on the brush, and only then, you can brush your teeth. It simply doesn't work the other way round! Now, can you place the sticker on the line for taking your toothbrush in your hand? And the next sticker, for putting the toothpaste on the toothbrush? And finally the last sticker for brushing your teeth?

A.2. Quantity stories

A.2.1. Books

Let's think about different amounts of things, for example, books. You could have only one book. Or the amount of books you can fit in your school satchel.

Or you could have a whole room full of books, like in a library. Can you imagine this? Let me place the first sticker on this line for one book (experimenter places sticker near one end of the line). This is the sticker for the medium amount, the satchel full of books; that's exactly here (experimenter places sticker near the middle of the line). Can you now put the sticker for the huge amount of books, for the whole library full of books, on the line?

A.2.2. *Candy*

When we think about different amounts of something, we can also think about something you surely like: candy! I could either give you only one candy. Or I could give you a whole handful of candy, or a whole bag full of candy. This is the sticker for only one candy, and I put it on the line exactly here (experimenter places sticker near one end of the line, in the same place as in the previous task). Now it's your turn. Where on this line would you put the sticker for the handful of candy? Can you put the third sticker on the line for the bag full of candy?

A.2.3. *Sand*

Have you played with sand in the sandbox on the playground? You can have different amounts of sand: you can have only one spoon full of sand, or you can have a whole bucket full of sand out of the sandbox, or you can even have a whole truck full of sand. Can you now place this sticker on the line for the spoonful of sand, the small amount, on the line? And now the sticker for the bucket full of sand? And the third sticker for the huge truck full of sand?

A.3. *Stories for integrating values*

Now I want to combine the two lines you have already learned (experimenter takes the last diagram with the three stickers on the vertical line and places three more stickers on the horizontal line as they appeared in the first task). I've got two little birds here (experimenter places toy birds on the first point along each line). Each of these birds can fly in only one direction — only straight ahead in the direction they are facing. They cannot turn — they can only fly in a straight line. Where the two paths of the birds cross is a meeting point. Look, I'll show you where the first two birds meet (experimenter shows how they move and the point where their paths meet). Let's place a sticker there (experimenter places blue at the meeting point). Now the birds start from these points (experimenter puts birds on next point of each axis). Can you place a sticker where they meet? You can also move the birds, if you like. And now they start from here (place birds on next point on each line), where do they meet? Now look: with these blue stickers we have combined time and amount. You remember, here (experimenter points at x -axis) was [time: breakfast time, lunchtime, dinnertime *or* the different amounts: one candy, a handful of candy, a bag full of candy]; and here

(experimenter points at y -axis), we put the stickers for [the different amounts: one candy, a handful of candy, a bag full of candy *or* time: breakfast time, lunchtime, dinnertime].

We can combine all this in a story: For example, imagine I'd give one candy to you after breakfast, that means, on our sheet, you see breakfast time (experimenter points to the sticker on appropriate axis) and one candy (experimenter points on the sticker on the appropriate axis); the blue point here combines both (experimenter points to blue sticker). After lunch, I give you a handful of candy; this is here (points to appropriate points on each axis and then the blue sticker that represents the integration of both values), and after dinner, I give you a whole bag full of candy (points to the relevant stickers).

We can connect the blue dots by drawing a line with a ruler here (experimenter points it with a pen). The blue line shows you how much candy you have over the day. In the morning you only have a bit, here, after lunch, you have some more, and at night before you go to bed, you have a whole bag full of it. I already drew a line like that one on our next sheet (produces next sheet). That line is just like the line we made wherever the birds met. This line (puts pen on the x -axis) stands for [time or amounts of things], and that line (experimenter puts pen on the y -axis) stands for [time or amounts of things]. This line here (experimenter puts pen on the graph) shows both time and amount combined. This line is a story. It's a story about how the amount changes as time goes by.

Now I'm going to tell you a story. Imagine you are sitting at home in the morning, having breakfast. There's a window next to the breakfast table, and out that window you can see the street in front of your house. A huge dump truck full of sand stops in front of your house, and very slowly, it starts to dump the sand next to the street. By the time you finish breakfast, there's a little amount of sand in front of your house, about as much as one spoonful of sand. When you look out of the window at lunchtime, it has dumped about as much sand as there is in a sandbox. And at dinnertime, the whole truckload of sand is in front of your house! Remember this line is our story. This sticker is for the amount of sand that the truck dumped by breakfast time. The day goes on, we can follow our line to see what happened. Here is the sticker for the amount of sand at lunchtime. Can you now place the sticker for the huge amount of sand the truck dumped by dinnertime?

Let's imagine another story. Do you like taking baths? I love taking baths! I like to sit in the bathtub and watch while the water fills up the bathtub. After about 1 minute, there is not yet much water in the tub. After about 5 minutes, the tub could be half full. And after 10 minutes, the bathtub is full of water. We can also show that on our story line. Look, here is 1 minute (experimenter shows a point on the appropriate axis) and here is a small amount of water (experimenter shows a point on the appropriate axis), and here is the combination of time and water on our story line (experimenter puts the sticker on the graph). Can you place the sticker for the amount of water after 5 minutes on our line? Now, can you put the last sticker for the full bathtub after 10 minutes on our story line?

A.4. *Stories for judgment tasks*

A.4.1. *Water*

Here we have two story lines. That means that a similar story happened twice. Imagine I was taking a bath at two different days; one day, it is one line, and the other day, it's the other line. Each line stands for filling up the bathtub on a different day. One day, I turned on the water full power, and the tub fills up much faster. Another day, I turned on the water just a little, so that the tub fills up more slowly.

A.4.2. *Sand*

These lines stand for two more stories. Remember the story I told you about the dump truck dumping sand in front of your house? Imagine it actually happened twice. Each line stands for dumping the sand on a different day. One day, the truck dumps the sand very fast. Another day, the truck dumps the sand very slowly.

A.5. *Judgment tasks*

A.5.1. *Rate judgment*

Look at this line (pointing at upper line). Does this line stand for the time it happened faster or for the time it happened slower?

A.5.2. *Quantity judgment*

That means that one time there was more and one time there was less. Look at this line (pointing at upper line). Does this line stand for the time when there was more or for the time when there was less?

Appendix B. Cover stories for Experiments 3 and 4

B.1. Age stories

I want you to think about age. For example, think about the different ages of people. In the beginning, a person is a baby. After that a person is a child. And after that a person is an adult.

B.1.1. People

I've got some stickers here. This sticker stands for a baby. I'm going to put this sticker here on the line for a baby (experimenter places sticker near one end of the line). So, this is a baby, okay? This is the sticker for a child, and I'm putting it right here (experimenter places sticker near the middle of the line). Now it's your turn: Can you put the sticker on the line for an adult? So, this is a baby, this is a child, and this is an adult (pointing at the correct stickers).

B.1.2. Dogs

Dogs are a lot like that. When a dog is born, it's a baby dog, and when it grows older, it's an adult dog, and then when it grows very old, it's an old, old dog. This sticker stands for a baby dog. Where would you put the sticker for a baby dog on the line? (experimenter gives the sticker to the child and the child places it on the line). This sticker stands for an adult dog. Can you put the sticker for the adult dog on the line? (experimenter gives sticker to child). This sticker stands for an old, old dog. Where would you like to put this sticker on the line? (experimenter gives sticker to child).

B.2. Rate stories

B.2.1. Vehicles

I want you to think about speed — how fast something goes. For example, let's think about vehicles. Vehicles go many different speeds. Some are very slow, like a bicycle, some are medium-speed, like a car, and some are very fast, like an airplane.

This sticker stands for something slow, like a bicycle. I'm going to put this sticker here on the line for something slow (experimenter places sticker near top end of the line for the downward mapping condition, or the lower end of the line for the upward mapping condition). So, this is for slow, okay? This is the sticker for something medium-speed, like a car. I'm putting it right here (experimenter places sticker near the middle of the line). This sticker stands for something fast, like an airplane. Now it's your turn: Can you put the sticker on the line for something fast? (experimenter gives sticker to child and child places sticker). So, this is slow, this is medium-speed, and this is fast (pointing at the correct stickers).

B.2.2. Dogs

Different kinds of dogs are like different kinds of vehicles — they run at all different speeds. A dog can be very slow, like a dachshund, or medium-speed, like a beagle, or very fast, like a German Shepard. This sticker stands for a slow dog, like a dachshund. Where would you put the sticker for a slow dog on the line? (experimenter gives the sticker to the child and the child places it on the line). This sticker stands for medium-speed dog, like a beagle. Can you put the sticker for the medium-speed dog on the line? (experimenter gives sticker to child). This sticker stands for really fast dog, like a German Shepard. Where would you like to put the sticker for a fast dog on the line? (experimenter gives sticker to child).

B.3. Size stories

B.3.1. Houses

I want you to think about size — how big something is. For example, let's think about houses. Houses are many different sizes. Some are very small, like a garden

house, some are medium-sized, like a farmhouse, and some are very big, like a palace.

This sticker stands for something small, like a garden house. I'm going to put this sticker here on the line for something small (experimenter places sticker near top end of the line for the downward mapping condition, or the lower end of the line for the upward mapping condition). So, this is for small, okay? This is the sticker for something medium-sized, like a farmhouse. I'm putting it right here (experimenter places sticker near the middle of the line). This sticker stands for something big, like a palace. Now it's your turn: Can you put the sticker on the line for something big? (experimenter gives sticker to child and child places sticker). So, this is small, this is medium-sized, and this is big (pointing at the correct stickers).

B.3.2. Dogs

Different kinds of dogs are like different kinds of houses — they come in all different sizes. A dog can be small, like a dachshund, or medium-sized, like a beagle, or very big, like a German Shepard. This sticker stands for a small dog, like a dachshund. Where would you put the sticker for a small dog on the line? (experimenter gives the sticker to the child and the child places it on the line). This sticker stands for medium-sized dog, like a beagle. Can you put the sticker for the medium-sized dog on the line? (experimenter gives sticker to child). This sticker stands for really big dog, like a German Shepard. Where would you like to put the sticker for a big dog on the line? (experimenter gives sticker to child).

B.4. Stories for judgment tasks

B.4.1. Rate judgment

Remember that this line (pointing to the appropriate line) tell us about age. This is where you placed the stickers for the different ages. And this line (pointing to the corresponding line) tells us about speed. This is where you placed the stickers for different speeds. When we combine age and speed we can make stories — and the last two lines left over are our story lines (pointing to the two function lines). These lines tell stories, and there are two of them. Those two stories are about two animals, and how they get faster as they get older. You've never seen these two animals before, and they don't look like any animal you've ever known. One is called a chimera and one is called an xyrus. They both get faster as they get older, but one of them does it more than the other. Look at this line (pointing at either upper or lower line). Does this line stand for the one that's faster or slower?

B.4.2. Size judgment

Remember that this line (pointing to the appropriate line) tell us about age. This is where you placed the stickers for the different ages. And this line (pointing to the corresponding line) tells us about size. This is where you placed the stickers for different sizes. When we combine age and size we can make stories — and

the last two lines left over are our story lines (pointing to the two function lines). These lines tell stories, and there are two of them. Those two stories are about two animals, and how they get bigger as they get older. You've never seen these two animals before, and they don't look like any animal you've ever known. One is called a chimera and one is called a xyrus. They both get bigger as they get older, but one of them does it more than the other. Look at this line (pointing at either upper or lower line). Does this line stand for the one that's bigger or smaller?

References

- Bates, E. (1993). Comprehension and production in early language development. *Monographs of the Society for Research in Child Development*, 58(3–4), 222–242.
- Bell, A., & Janvier, C. (1981). The interpretation of graphs representing situations. *For the Learning of Mathematics*, 2, 34–42.
- Bryant, P. E., & Somerville, S. C. (1986). The spatial demands of graphs. *British Journal of Psychology*, 77, 187–197.
- Gattis, M. (2001). Mapping conceptual and spatial schemas. In M. Gattis (Ed.), *Spatial schemas in abstract thought* (pp. 223–245). Cambridge, MA: MIT Press.
- Gattis, M. (2002). *Mapping relational structure in spatial reasoning*. Manuscript under review.
- Gattis, M., & Holyoak, K. J. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(1), 231–239.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155–170.
- Gentner, D. (1988). Metaphor as structure-mapping: The relational shift. *Child Development*, 59, 47–59.
- Halford, G. S. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Handel, S., De Soto, C. B., & London, M. (1968). Reasoning and spatial representations. *Journal of Verbal Learning and Verbal Behavior*, 7, 351–357.
- Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, 67, 2797–2822.
- Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words? *Cognitive Science*, 11, 65–100.
- Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research*, 60, 1–64.
- McDermott, L., Rosenquist, M., & vanZee, E. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, 55, 503–513.
- Piaget, J., Grize, J., Szeminska, A., & Bang, A. (1968/1977). *Epistemology and psychology of functions*. Boston: Reidel.
- Savage-Rumbaugh, E. S. (1993). How does evolution design a brain capable of learning language? *Monographs of the Society for Research in Child Development*, 58(3/4), 243–252.
- Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.
- Tversky, B. (1995). Cognitive origins of graphic conventions. In F. T. Marchese (Ed.), *Understanding images* (pp. 29–53). New York: Springer.
- Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology*, 23, 515–557.
- Wilkening, F. (1981). Integrating velocity time and distance information: A developmental study. *Cognitive Psychology*, 13, 231–247.