Modeling the Logarithmic-To-Linear Shift In Representations Of Numerical Magnitudes

Leif K. Lende (leif.k.lende@diakonhjemmet.no)
Department of Social Education and Occupational Therapy, Diakonhjemmet University College Rogaland
Vaagsgaten 40, N-4306 Sandnes, Norway

Niels Taatgen (n.a.taatgen@rug.nl)
Department of Artificial Intelligence, University of Groningen
Nijenborgh 9, 9747 AG Groningen, Netherlands

Abstract

We present findings from an experiment with 6-year-old children in Norway who are getting started with the process of learning the numbers from 1 to 10. The findings provide information on number line estimation - that is, translating a number to a spatial position on a number line. The results show different categories of representation of the magnitudes on the number line, which may represent different stages in a learning sequence. On this basis, we show a proposal for a cognitive model of the learning process towards a linear representation of magnitudes.

Keywords: Learning; numerical magnitudes; number line; dynamic decision making; memory; cognitive architectures; ACT-R.

Introduction

In this paper we will present our first findings from the pre-test of a larger experiment where the subjects are Norwegian 6 year old children just started on pre-school education learning numbers from 1 to 10. One important learning process that is involved when dealing with number magnitude is the estimation of what position a number value has on a number line.

The learning sequence involved is the one that Siegler calls the logarithmic-to-linear shift in representations of numerical magnitude (Siegler, Thompson, & Opfer, 2009). In order to make a cognitive model of this learning process we want to investigate what cognitive processes are involved in learning the positioning of number magnitudes on a number line.

Siegler et al (2009) show that children undergo parallel changes from logarithmic to linear representation on numerosity estimation tasks. The example we have reused from their article in figure 1 shows long-term changes in estimation of whole number magnitudes. (A) On 0–100 number lines, kindergartners' estimates were better fit by the logarithmic function than by the linear, whereas second-graders' estimates were better fit by the linear function than by the logarithmic; (B) On 0–1000 number lines, second-graders' estimates were better fit by the logarithmic function than by the linear, whereas fourth-graders' estimates were better fit by the linear function than by the logarithmic.

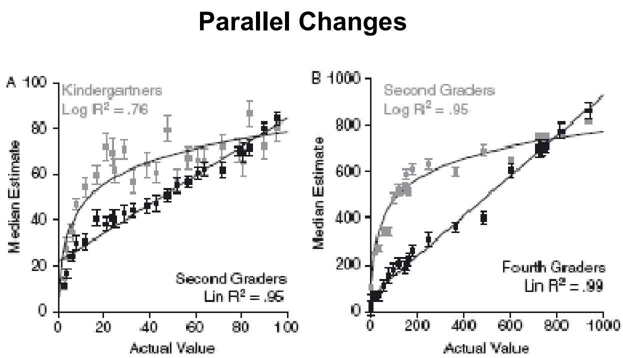

Figure 1. The logarithmic to linear shift. From Siegler, Thompson, & Opfer, (2009), Copyright 2009 Wiley. Reprinted with permission.

From Siegler et al's research and from others (Dehaene, Izard, Spelke, & Pica, 2008) we know that a logarithmic representation is likely to be the starting point of this learning sequence. This representation is usually attributed to Weber's law. However, in some of the Ramani and Siegler's work (2008) the initial performance is not logarithmic, but on average linear but incorrect. Maybe this has something to do with the scale: 1..10 as opposed to 1..100 compared to the size of the task sheets. But then the simple Weber explanation will not hold universally.

We also know from the research of Siegler and colleagues (2009) that this learning sequence happens several times at different stages during lifetime. That means children can have multiple representations of numerical magnitudes. As an example they can have a logarithmic representation of the numerical magnitudes on a number line from 0 to 100 while they have a linear representation on the number line from 0 to 10. As the children get older and learn a linear representation on the number line from 0 to 100 too, they may still have a logarithmic representation on the number line from 0 to 1000, and so on.

A possible account for the transition from logarithmic to linear is that children learn the location of particular points on the number line. Schneider et al. (2008) have investigated the role of anchor points in an eye-movement study. They confirmed in a cross-sectional design with children from Grades 1 to 3 that the distance from the nearest orientation point is the most important and highly significant predictor of the frequency with which a position on the number line is fixated. Furthermore the results show that the distribution of fixations on the number line for all three groups of first grade, second grade and third grade children are concentrated around beginning, midpoint and ending of the number-line, see figure 2.

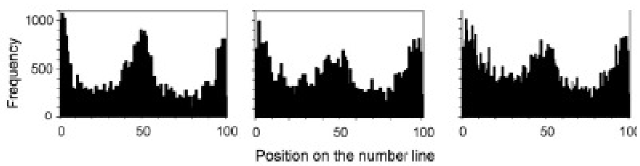

Figure 2: Distribution of fixations on the number line (left: first grade; middle: second grade; right: third grade). From Schneider et al. (2008), Copyright 2008 Elsevier. Reprinted with permission.

This shows that a great deal of attention is directed to those three positions on the number line and that those positions are of importance when a subject is doing a number line estimation task. As Schneider himself notes, the results then replicate Petitto's (1990) findings with respect to the use of the midpoint strategy and a counting-up strategy by students in Grades 1–3.

By an anchor point on a number line we mean a location that the positions of other number values are related to. If zero is an anchor point we can say: I know where zero is, and I know that two is a bit to the right of zero's position, and then the same for all of the actual distribution of number values. That means there is a mapping going on between a number value and a position on the number line that is related to an anchor point. The work of Schneider et al (2007) suggests at least three such anchor points: the endpoints and the midpoint. They also show that from grade 1 to3 children tend to increasingly focus on the correct positions on the number line while solving the estimation tasks.

A possible model of progressing towards a linear time scale can therefore be one that increasingly learns the locations of particular points on the number line, and uses those as anchors to determine the points that it does not know. It therefore needs some sort of representation of the positions of anchor points, but also a method of determining points in between those anchors.

As a theory of how anchor points are stored in memory, we use the ACT-R declarative memory (Anderson, 2007). In order to determine positions between the anchor points, we use the optional mechanism of *blending* (Lebiere, Gonzalez, & Martin, 2007). This mechanism can generate weighted averages over examples in memory based on their activation and the distance from the requested item.

Lebiere and colleagues (2007) used blending in an instance-based model of decision-making for repeated binary choice. An example of blending that is more similar to numeric magnitudes is one of time perception by Taatgen & van Rijn, (Submitted).

To investigate the usefulness of this idea, we looked for evidence of anchoring in collected data of our own.

The number line estimation task

The experiment, from which we will use the findings from the pretest in this paper, is a sort of replication of a study first made by Siegler and Ramani (2008) among preschool children from low income families. In Norway the population of children at kindergarten and preschool are mixed. Thus we have defined the population for the experiment by learning level. In Norway children start at school the year they are 6 years old. This first year in school they start to learn the numbers from 1 to 10. We assume that this represents a mental level that should make most of them just capable of dealing with the empty number line.

The Method of the pre-test

Participants

Participants were 39 children in their first year at school, so called preschool, with no experience with number lines. All of them are born in 2004 and recruited from the same municipality, Gjesdal. 17 of them are recruited from Solås School, 7 from Dirdal School and 15 from Bærland School. The population at these schools is mixed, but at Bærland with a larger representation of children with two languages, Norwegian not being their mother tongue.

Materials

Stimuli for the number line estimation task were two stacks of 10 sheets of paper, each with a 25 cm long line arranged horizontally across the page, with '0' just below the left end of the line, and '10' just below the right end. A number from 1 to 10 inclusive was printed approximately 3 cm above the center of the line, with each number printed on one of the 10 sheets in each stack. The order of the sheets in the stack is randomized.

Procedure

The test is conducted as a teacher to student task:

- The teacher or student pulls a sheet from the stack.
- The teacher says: "Here's the number [number that is on the pulled sheet]. And here you see a line that starts with 0 and ends at 10. Where on this line is the correct position for the number you see. Put a mark with your pencil".
- The student makes a mark where he thinks the number should be positioned.

In this way the task is carried out with all the sheets in the first stack. Then the task is continued in the same way with the second stack. In this way the numbers from 1 to 10 inclusive were presented twice in random order, with all numbers presented once before any number was presented twice. No feedback was given, only general praise and encouragement.

Result and discussion

Figure 3 shows the mapping between numbers and positions on the number line that we found in the experiment. Performance is on average reasonably good. What is surprising is that the extent in which the curve differs from linear is not towards a logarithmic curve, but in the opposite direction. Points are plotted with error bars.

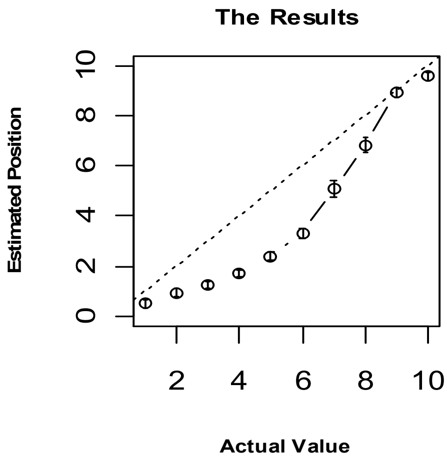

Figure 3. Mean task result for 39 children

To get a better sense of the difference of responses, we categorized individuals into four categories. From this categorization, three of the participants were excluded because they produced more or less random results.

We found that there were roughly four groups that resemble a horizontal mirrored L, a mirrored Z, an S and a linear line. Even though the borders between these categories are fuzzy, they seem to represent a sequence in learning the number line from 0 to 10. The categories are classified as follows:

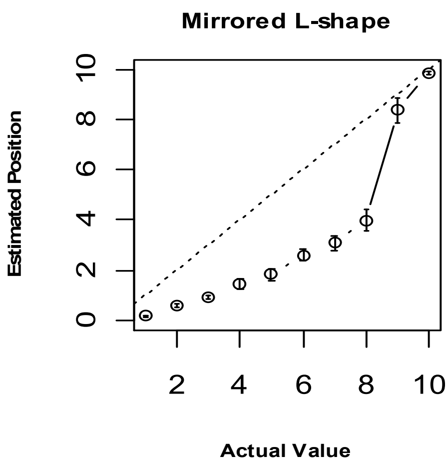

Figure 5. Mean of results from category L-shape

The mirrored L is represented by subjects that seem to position mostly by counting up and know something about one or two numbers close to zero and ten. They are classified in this category if 1 through 8 are all far below the midline, and 10 above. The plotting of the mean values gives something like a vertical mirrored L, see figure 5.

The next group, from whom the plotting of mean values of the task result give us a vertical mirrored Z image, has positioned more of the numbers close to ten on the correct side of the middle of the number line and are classified in this group if at least 8-10 are far above five and 1-6 far below:

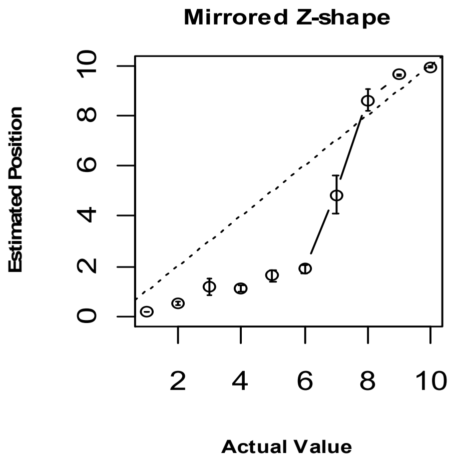

Figure 6: Mean of results from category Z-shape

The third group is even more close to a linear representation and the plotting of the mean values of this group gives us an image more like an italic S. The subjects in this group are classified to this category if 5 or 6 are put around the midpoint, and 1-4 are below, and 7-10 are above the line. The new learning step for this group is that they have obviously recognized one or two number magnitudes close to five to belong to the middle of the empty number line:

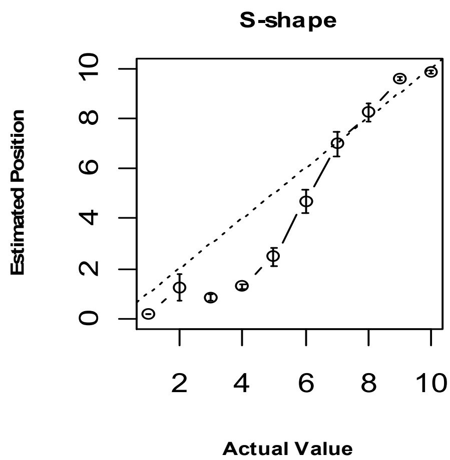

Figure 7. Mean of results from category S-shape

The fourth and last group is the subjects that seem to have grasped it all. The plotting of mean values of the task result from this group looks linear and gives us an image close to a linear graph:

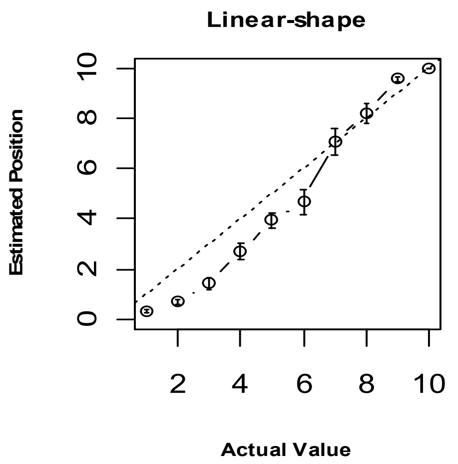

Figure 8: Mean of results from category Linear-shape

Our hypothesis is that this is also the order of progress in the involved learning sequence: mL -> mZ -> S -> linear.

A computational cognitive model

The results from the eye-tracking experiment of Schneider and colleagues, mentioned above, makes us believe that cognitive functions dealing with number line estimation connects to one or more anchor points. Thus, a general idea of this model is that it stores a set of anchor points of which it knows the location of on the line. These anchors are stored as chunks in declarative memory, more specifically as a pair of number and position. When the model wants to put a point on the line, it attempts a retrieval of that number and its position.

How the model works

In order to allow the model to mark positions of numbers for which it does not have a memory representation, we will use the blending mechanism that we mentioned earlier. Blending uses existing (anchor) chunks in declarative memory to construct intermediate representations. Two factors are important in this construction process: the mismatch between the anchor point in memory and the number we look for, and the memory strength of the anchor point.

An appropriate function to define the mismatch between two numbers is Weber's law. Weber's law is typically used to express a "just noticeable difference" between two stimuli. We use it to calculate the magnitude of the perceived distance between values like this:

$$Dist = Weber's\ constant * \frac{abs(num2 - num1)}{max(num2, num1)}$$

This perceived distance can then be used as the mismatch in the activation equation. Weber’s constant becomes a scaling parameter.

We think that a subject that is involved in the learning sequence we are investigating, knows some anchor points and has a strategy for relating the remaining number values to those points.

The activation of an anchor point is equal to its base level activation that is based on the number of experiences with that anchor in the past minus the distance (according to the Weber formula) between it and the number you are trying to retrieve. By first using the equation for base level learning of ACT-R the model uses the softmax equation to determine P_i, the probability of retrieving each qualifying chunk.

$$P_i = \frac{e^{A_i/t}}{\sum_j e^{A_j/t}}$$

.

In this equation t is a constant for the retrieval noise level and is set to its default level of 0.25. For the blending retrieval our model uses the value of P_j to calculate the result value that is retrieved by the formula of Taatgen and van Rijn (2010):

$$Result\ value = \sum_j P_j V_j$$

Vj is the position of the anchor point on the number line. So, the resulting position for the number on the number line is a weighted average of the locations of anchors (weighed according to their retrieval probability).In our model the case is not time intervals measured by number of pulses, but positions (where is the number on the number line), that are stored in a memory pool.

The functions of the model

The model is a very simple model, a general memory theory. And it is not the complete story. Counting will be implemented at a later stage. Now it has an activation baseline function and there are three functions dealing with the declarative memory.

One function makes a reference list of numbers involved and their position on the number line. A chunk is represented as a list, with a number (what number is it about) and a position (where is it on the number line), and a references list with moments in time the chunk has been accessed.

The mismatch function is based on Weber’s law and the result value is zero, a negative value or a positive value depending on whether the first number is similar, less than or greater than the second number. The mismatch assumes two numbers are more similar if they are closer and higher and is used to calculate the activation of a chunk.

A blending function is performing the retrieval and adds noise. Because of that we do not use the regular ACT-R retrieval rule and noise activation function.

The blending retrieval

By using the four functions mentioned above, the model simulates one of each of the four cases involved in our empirical data. We have the L-shape case, where we assume that we know one better than ten. Then chunk containing number and position for the anchor point 1 is given one more activation entry. In the second case, the Z-case, we assume that we know 10 a bit better. In the third case, the S-case, we assume that we now also know five. And in the fifth case, the linear case, we know pretty well where most or all numbers belongs on the number line. The number of entries we add to the chunk can also be used to manipulate the simulation of the model. The following table shows the number of activation entries used for the anchor points:

Table 1: Number of activation entries

Anchor points:	**1**	**3**	**5**	**8**	**10**
Logarithmic	1				1
Mirrored L-shape	2				1
Mirrored Z-shape	5				3
S-shape	6		1		4
Linear shape	7	1	4	1	5

Results from running the model and discussion

The model gives five images of simulated values where four are related and rather similar images of estimated values to images from mean values of our four groups of empirical data, when plotted.

The initial stage, is a case that gives a logarithmic curve and has the smallest activation of only one entry on each of 1 and 10 as anchor points, see figure 9.

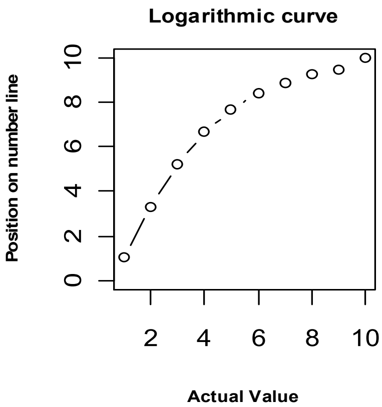

Figure 9: Model result from smallest activation

The next case tries to simulate the mirrored L-shape category. In this case the anchor point 1 is given one more activation entry. The image we get from the curve of this case is not so clearly L-shaped, figure 10. The reason for this we assume is because there is a great bit of counting involved in children's estimation at this level. Our model does not do counting as we rely solely on Weber's law and blending.

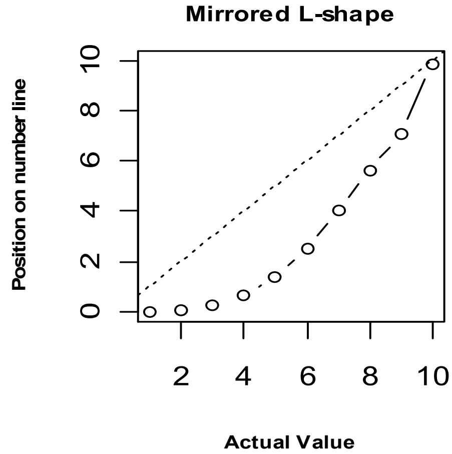

Figure 10: Model result simulating L-shape group

In the next case we try to simulate the category when the results give a curve like a mirrored Z-shape. Now we know one and ten a bit better, and the activation in declarative memory is set to five entries for the anchor point 1 and three entries for anchor point 10. As figure 11 compared to the curve image in figure 6 shows, the model now fit a lot better.

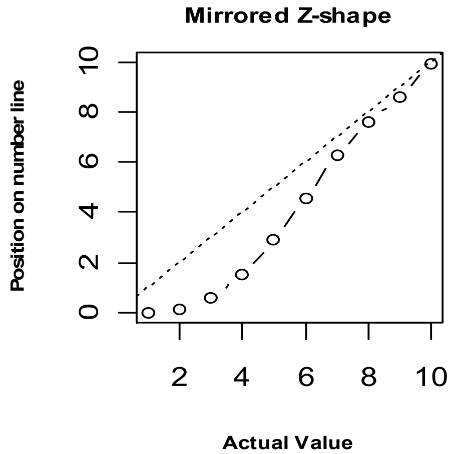

Figure 11: Model result simulating Z-shape group

In the case where we get an S-shaped curve the model knows the anchor point 5 too, but with smallest activation, and 1 and 10 is strengthened with one entry each. We now clearly see that the image of the curve is more and more similar to the relevant curve image from the result of our experiment data. The reason for that we assume is relaying on the subjects decreasing use of counting as they learn more anchor points and therefore corresponds better to our model as we already has mentioned do not use counting but only blending, see figure 12.

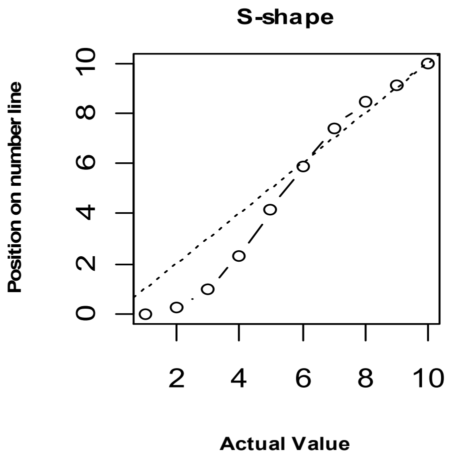

Figure 12: Model result simulating S-shape group

In our last case, the model knows even more anchor points. We have added 3 and 8 as new anchor points in declarative memory, but with smallest activation: one entry. Again the anchor points 1 and 10 are strengthened by one more entry. The anchor point 5 is strengthened by three more entries and has now almost the same activation level as the anchor point 10.

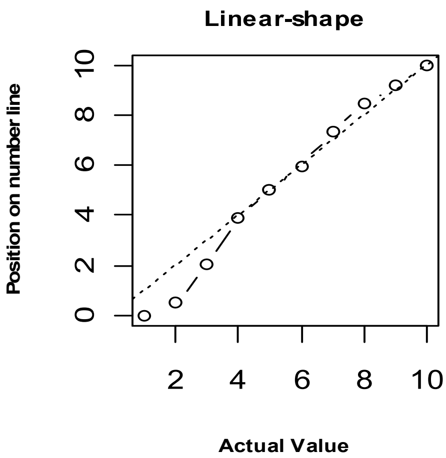

Figure 13: Model simulating Linear-shape group

In our data it seems like all of the subjects that understand the task use counting as an important part of the strategies for estimation. As we can see, in the same way as the results of our collected data from 6 year old children showed, we obtain no logarithmic curve from running our model. If we investigate the physical size of the unit used by the subjects in counting up or down from an anchor point, it is for all of them much smaller than a tenth of 25 cm, that was the length of the number line used in the estimation task. That shows that for the counting strategy the subjects do not have a clear clue of what the size of a unit should be.

But on the opposite, with a larger scale, for example up to 100, the children's unit will be too large and counting often will produce a logarithmic curve like Siegler and others has found.

Conclusion

The model produces a rather good fit to our data from real life. If counting-up and counting-down were added to the model's simulation we assume that it would give an even better fit to our data. To adapt it to estimation tasks with number lines of longer length then 0 to 10, there is a need of implementing scaling to simulate the physical size of the number line.

Acknowledgments

Thanks to Gjesdal municipality for permission and even funding of our experiment, given by teachers use extra time paid by the municipality to assist in our experiment.

References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* (p. 290). Oxford University Press. Retrieved February 1, 2011

Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science*, *320*, 1217-20

Lebiere, C., Gonzalez, C., & Martin, M. (2007). Instance-Based Decision Making Model of Repeated Binary Choice. *Proceedings of the 8th International Conference on Cognitive Modeling* (pp. 67-72). Oxford UK: Taylor & Francis.

Petitto, A. L. (1990). Development of Numberline and Measurement Concepts. *Cognition & Instruction*, (1), 55

Schneider, M., Heine, a, Thaler, V., Torbeyns, J., Desmedt, B., Verschaffel, L., et al. (2008). A validation of eye movements as a measure of elementary school children's developing number sense. *Cognitive Development*, *23*(3), 409-422.

Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental science*, *11*(5), 655-61.

Siegler, R. S., Thompson, C. a, & Opfer, J. E. (2009). The Logarithmic-To-Linear Shift: One Learning Sequence, Many Tasks, Many Time Scales. *Mind, Brain, and Education*, *3*(3), 143-150.

Taatgen, N., & van Rijn, H. (Submitted). Traces of Times Past: Representations of Temporal Intervals in Memory.