Application of the Category Adjustment Model in Temporal, Spatial, and Abstract Magnitude at the Billions Scale

Ilyse Resnick (ilyse.resnick@temple.edu)
Department of Psychology, 1701 N. 13th Street
Philadelphia, PA 19122 USA

Thomas F. Shipley (tshipley@temple.edu)
Department of Psychology, 1701 N. 13th Street
Philadelphia, PA 19122 USA

Abstract

The current study examines the generalization of the Category Adjustment Model (CAM) across scales along two dimensions: time and distance. Participants were presented with geologic time and astronomical distance information either conventionally or using the hierarchical alignment model. Participants provided with hierarchically structured magnitude information for time and distances were more accurate on similar estimations at large scales than participants given the same content in a conventional manner. Patterns in event and distance estimation, along with overall group differences, are consistent with the CAM; suggesting people use hierarchically organized categorical information when estimating across scales and dimensions, and providing salient category boundary information improves estimation. Findings suggest a common representation of scale information for temporal, spatial, and abstract (numeric) magnitudes. Patterns of abstract magnitude estimations are consistent with segmented linear models of scale representation. Implications of the CAM in scale representation and the hierarchical alignment model in education are discussed.

Keywords: Category Adjustment Model; Hierarchical Alignment; Scale Representation

Introduction

The Category Adjustment Model (CAM) is an adaptive Bayesian account for the pattern of systematic biases observed in recall of metric quantities due to category membership (Huttenlocher, Hedges, & Prohaska, 1988; Huttenlocher, Hedges, Vevea, 2000). The CAM posits 1D, 2D, and 3D magnitudes are stored in a hierarchical combination of metric and categorical information. In the absence of lower-level information (e.g., precise metric information), people use higher-level categories to aide in estimation. Variation in estimation, therefore, occurs due to imprecision of category boundaries. Recall is biased towards the 'prototype' of the respective category. For example, when recalling the position of an object in a circular display, participants naturally divide the circle into mental quadrants and the recalled location is biased towards the center (or prototype) of the relevant quadrant (Huttenlocher, Hedges, & Duncan, 1991).

The CAM predicts recall patterns on a range of dimensions (e.g., fatness of fish, grayness of squares, and lengths of lines (Huttenlocher, et al., 2000), events (Huttenlocher, et al., 1988), and even social dimensions such as perception of facial expressions (Roberson, Damjanovic, & Pilling, 2007) and judgments of gender and ethnicity (Huart, Corneille, & Becquart, 2005)). However, there is limited research examining the CAM's predictive capability for a given dimension (such as temporal and spatial scales) across different scales (such as from human scales through to scales outside of human perception). Science education research has identified conceptual categories for spatial and temporal scales outside of human perception (e.g., Trend, 2001; Tretter, Jones, Andre, Negishi, & Minogue, 2006), suggesting people may conceptualize magnitude information at relatively small and large temporal and spatial scales using a combination of metric and categorical information. Resnick, et al. (2012) experimentally assessed the role of categories in estimations of large temporal magnitudes. Participants who were provided with salient hierarchically organized event boundaries fostered a linear representation of events on the Geologic Time Scale compared to those who received the same information about the events without the salient hierarchical structure. Aligned with the CAM, this finding suggests the use of hierarchically organized category boundaries in the representation of events at larger temporal scales.

The current study aims to add to this relatively sparse literature by examining the generalization of the CAM across scales and dimensions. Two main objectives are to replicate research on memory for large temporal magnitudes (geologic time), and extend research to another dimension: space. Astronomical distance (a spatial magnitude at a large scale) was chosen for two reasons. There is already extensive research on CAM and spatial distance; demonstrating spatial distances at familiar scales are stored in a combination of metric and categorical information (e.g., Huttenlocher, et al., 1991; Huttenlocher, et al., 2000). Additionally, while the precise nature of the relationship is unclear, there is a systematic relationship between time and distance (e.g., Clark, 1973; Gentner, 2001), suggesting that time and distance at human scales are represented and estimated in the same way. Thus, if temporal and spatial dimensions across familiar and relatively larger scales are represented in a similar way, an analogous pattern of memory performance would be expected.

Relevant to the current study, the CAM makes two predictions. First, estimations of temporal and spatial magnitude should be biased towards the prototype of each event or object's category. There is evidence that suggests people with a moderate amount of knowledge regarding geologic time (e.g., in-service science teachers), divide the Geologic Time Scale (4.6 billion years) into three categories: 'extremely ancient', 'less ancient', and 'geologically recent' (Trend, 2001). It is beyond the scope of this paper to identify and characterize the types of categories used by novices to represent large temporal and spatial magnitudes. Rather, the current study will assess if providing salient internal structure of magnitude relations, through the use of the hierarchical alignment activity (Resnick, et al., 2012), improves estimation of large temporal and spatial magnitudes. In this way, the current study examines a second prediction of the CAM: people with salient internal structure of magnitude relations within hierarchically organized category boundaries should have more linear representations of magnitude compared to those who do not.

The current study also examines patterns of abstract (numeric) magnitude estimation (i.e., not content-specific) at the same scale as geologic time and astronomical distance. One common property of time and distance is they are both one-dimensional vectors (e.g., Clark, 1973; Gentner, 2001), as is abstract magnitude. Similar patterns in overestimation of small magnitudes and underestimation of large magnitudes are found with estimations of geologic events (Libarkin, Kurdziel, & Anderson, 2007), astronomical distance (Miller & Brewer, 2010), and abstract magnitude (Siegler & Opfer, 2003). Studies of abstract magnitude suggest this pattern of errors may be due to compressive effects of unfamiliar magnitudes on a mental number line (see Barth & Paladino, 2011 and Opfer, Siegler & Young, 2011 for discussion of competing models). Consistent with the scale of geologic time and astronomical distances, the current study will examine abstract magnitude at two scales: million and billion. Number word frequency studies suggest that there may be differences in the representation of the million and billion scales, because the frequency of occurrence influences the structure of representation and the number 'million' appears more frequently than 'billion' (e.g., Dehaene & Mehler, 1992). Thus, sampling from across the million and billion scales may reveal potential representational differences between the two scales.

While research has not explicitly examined the CAM in abstract (numeric) magnitude representation, there are a number of studies that look at the role of the subjective categorization of numbers in estimation (e.g., Laski & Siegler, 2007; Mix, Huttenlocher, & Levine, 2002; Siegler & Robinson, 1982). Findings suggest that individual numbers can serve both as their own distinct category (a specific quantity of something) as well as part of a set of numbers (e.g., 'small' versus 'big' numbers) (Mix, et al., 2002). Further, children who spread numbers evenly across group dimensions were more accurate on an abstract magnitude task than those who grouped more numbers into one 'big' category (Siegler & Robinson, 1982). The current study will examine if the presentation of salient category boundaries in specific dimensions transfers to abstract magnitude representation. Because participants will be working with magnitudes with temporal and spatial content, transfer to abstract magnitude should occur. If the CAM accounts for abstract magnitude at large scales, similar patterns of estimation are expected for geologic time, astronomical distance, and abstract magnitude.

Methods

Participants

Forty participants were recruited from an undergraduate psychology experiment pool (20 in the hierarchical (experimental) group and 20 in the conventional (control) group). The demographics of the participants were consistent with a large urban American university.

Hierarchical Design In the hierarchical alignment condition, participants completed the same hierarchical alignment activity developed by Resnick and colleagues (2012), which is based on the progressive alignment model (Kotovsky and Gentner, 1996; Thompson & Opfer, 2010). Participants made ten separate time lines, aligning time to a horizontal one meter space. They began with a familiar personal time scale, working through different historic and geologic time lines, up to the full Geologic Time Scale. For each time line, participants were given a partially completed time line, and were required to label the time line's length (in years) and locate where all previous time lines would begin on the current time line (see Figure 1).

Hierarchical organization highlights how each temporal scale is related to the other scales. Practice mapping magnitude relations across scales provides internal structure of magnitude relations within each scale. Thus, the hierarchical organization helps to populate each scale with additional categorical boundary information.

The current study developed a new analog version of the temporal hierarchical alignment activity for spatial distances (see Table 1 and Figure 1). For the hierarchical alignment of spatial distances, participants align ten increasingly larger scales of distance to a one meter space, beginning with a familiar distance. The hierarchical alignment condition takes approximately 45 minutes to complete.

Conventional Design The study sought to contrast the intervention with a realistic training program similar to one that might be used to instruct students in a classroom on these scales. Common pedagogical approaches to teaching geologic time (Libarkin, et al., 2007) and astronomical distances (Miller & Brewer, 2010) are to create spatial analogies, such as placing events or objects in the correct sequence. Participants completed ten separate puzzles, placing the events/objects into the correct sequence. The puzzles were made up of pieces of paper, half containing magnitude information and half with the respective category information. Participants were required to match the magnitude information with the corresponding category information for each scale, and place the scales in the correct sequence. The first puzzle represented the first temporal/spatial scale (see Table 1), with each puzzle representing an increased amount of magnitude. The tenth and final puzzle represented all of geologic time/distance to Makemake. The conventional condition took approximately 45 minutes to complete.

The conventional and hierarchical conditions were aligned on the following properties: number of scales, number of times participant identifies each scale (i.e., the first scale is identified ten times; the last scale is identified once), progressive increase of magnitude, information provided about each event/object, and total length of time on task. Thus, the only difference between conditions was the hierarchical alignment of scale information.

One potential difference between the temporal and spatial information was identified. Participants are likely to be familiar with thinking about temporal scales extending back hundreds of years ago; learning about recent human history is common. However, participants may not have the same level of familiarity with conceptualizing the vertical nature of the spatial scales. Because it is likely people have more experience traveling parallel to Earth's surface, or 'horizontally', as opposed to traveling vertically away from Earth's surface, we used this horizontal experience as an initial introduction of the vertical scale. As a way to familiarize participants with the vertical scale, a horizontal map was presented for each of the first three scales in both the hierarchical and conventional conditions. The maps showed an eleven, fifty-two, and four-hundred mile radius extending out from the university where the study took place. To engage the participants in grounding this scale to their personal experience, participants were asked if they had been anywhere on that radius or if they were familiar with the area. Because participants likely do not have experience thinking about larger temporal scales, no map was provided for the remainder of the spatial scales.

Table 1. List of Temporal and Spatial Scales, including category names and magnitude information

Temporal Scale	Years	Spatial Scale	Miles
Personal	20	Troposphere	11
Human Lifespan	75	Middle Atmosphere	52
American History	519	Exosphere	400
Recorded History	5,512	Inner Van Allen Radiation Belt	6,000
Human Evolution	6,000,000	3753 Cruithne (quasi-satellite)	8,450,000
Cenozoic	65,000,000	Mercury	57,000,000
Phanerozoic	542,000,000	Saturn	777,000,000
Proterozoic	2,500,000,000	Neptune	2,700,000,000
Archean	3,800,000,000	Pluto	3,580,000,000
Hadean	4,600,000,000	Makemake (dwarf planet)	4,800,000,000

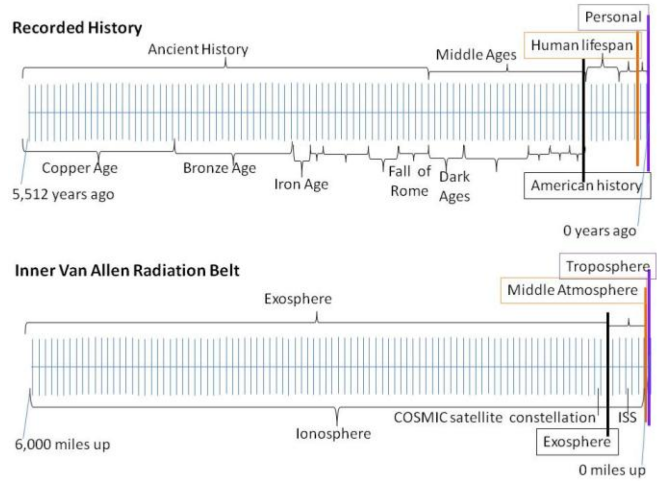

Figure 1. Example of a temporal and spatial number line at the thousands scale in the hierarchical condition. Note: the three previous temporal and spatial number lines are located relative to the current scale.

Procedure In a two-hour session, participants were presented with information about time and distance, with both presented as either hierarchically or conventionally (~90 minutes). Participants across conditions then completed the same assessment measures (~30 minutes).

Measures A series of line estimation tasks were developed to assess participants' representations of geologic time, astronomical distances, and abstract (numeric) magnitude. Line estimation tasks are commonly used to assess mental scaling of abstract magnitude (e.g., Ebersbach, et al., 2008; Siegler & Booth, 2004).

To measure representation of events on the Geologic Time Scale, an item from the Geoscience Concept Inventory (GCI), a reliable and valid instrument measuring a range of geoscience knowledge (Libarkin, et al., 2005), was adapted as a number line task. The GCI item presents participants with five time lines, with the following four geologic events placed in different locations: life appears, dinosaurs appear, dinosaurs disappear, and humans appear. Participants are required to choose the correct linear representation, with the other four time lines representing common misconceptions. In order to capture more variance in participants' representations, the GCI item was adapted so that participants were given a blank time line (anchored by 'present day' and 'Earth forms'), and asked to locate the same four events as used in the GCI item.

To measure representation of objects on an astronomical scale, an item was developed as an analog to the geologic event time line described above. Here, participants were presented with a blank number line (anchored by 'Earth's surface' and 'Makemake'), and asked to locate four objects on the same scale as on the event time line: Pluto, Mars, Mercury, and Cruithne.

To measure representation of abstract (numeric) magnitude (not content specific) a series of line estimation tasks were given. Participants were given a sentence stating when/where an event/object was, and then asked to locate that magnitude on the number line (e.g., "Venus is 26 million miles away from Earth. Please draw on the line provided where Venus is located."). These items were framed in terms of objects and events to match the form of the other experimental measures. These estimations are considered estimations of abstract magnitude because the participants are explicitly given a magnitude to place on the number line; no recall is required. The questions provide the numerical values and ask for an estimation of the appropriate location on a spatial scale. To assess representations of the millions and billions scale, participants were asked to estimate two 'events' and two 'objects' on a 4.6 billion scale, and two 'events' and two 'objects' on a 542 million scale.

Results

Participants in the hierarchical condition were more accurate overall on the event time line estimation task ($t(38)=2.67$, $p=.01$) and the object distance task ($t(38)=3.02$, $p=.01$) compared with participants from the conventional condition. On both tasks, this effect is driven primarily by the estimation of the 2nd and 3rd events/objects. Participants across conditions performed similarly when placing the 1st (life appears/Pluto) and 4th (humans appear/Cruithne) events/objects on the number line ($p>.05$). However, participants in the hierarchical condition were significantly more accurate when placing the 2nd (dinosaurs appear) ($t(38)=2.79$, $p=.01$) and 3rd (dinosaurs disappear) ($t(38)=2.53$, $p=.02$) events on the time line, and the 2nd (Mars) ($t(38)=3.38$, $p<.01$) and 3rd (Mercury) ($t(38)=2.79$, $p=.01$) objects on the number line compared to the conventional condition (see Figure 2).

Performance across groups on the object distance estimation task was significantly more accurate than on the event time line estimation task ($t(39)=2.85$, $p=.01$).

The eight abstract (numeric) magnitude line estimation tasks were highly correlated ($rs > .529$, $p<.01$) and had strong internal consistency (Cronbach's alpha=.94). There was no difference in performance when estimating abstract magnitude when estimations were temporally framed compared with spatially framed ($p>.05$). Given the high correlations, strong internal consistency, and no performance differences between items that were temporally and spatially framed; a single abstract magnitude scale was created. Participants from the hierarchical condition were significantly more accurate on the abstract magnitude scale (μ error = 11.50mm) than the conventional condition (μ error = 30.14mm) ($t(25.38)=2.58$, $p=.02$). That the participants from the hierarchical condition are more accurate on the abstract magnitude scale than participants from the conventional condition is consistent for estimations on both the million and billion scales. Across conditions, participants were significantly more accurate when making estimations on the millions scale (μ error = 14.73mm) compared with estimations on the billions scale (μ error = 26mm) ($t(39)=3.45$, $p<.001$).

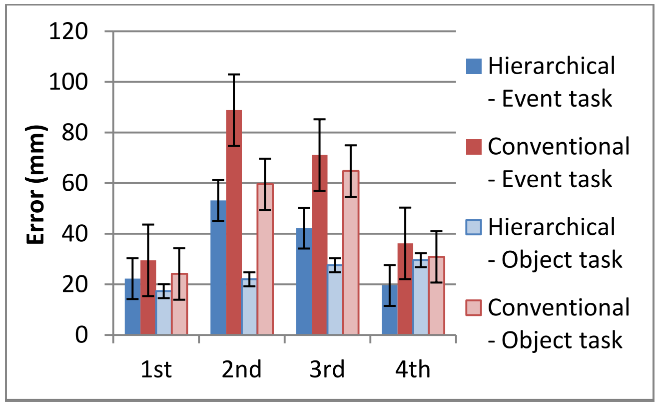

Figure 2. Average error (mm) for hierarchical and conventional conditions on the event and object line estimation tasks. For the event/object line estimation tasks, 1st = Life/Pluto, 2nd = Dinosaurs appear/Mars, 3rd = Dinosaurs disappear/Mercury, and 4th = Humans appear/Cruithne, respectively.

Discussion

The current study successfully replicated the Resnick, et al. (2012) findings; participants provided with hierarchically structured event information were more accurate on event time line estimations than participants given the same content in a conventional manner. Here we found a similar result for astronomical distances and abstract (numeric) magnitude. These findings are aligned with the CAM, suggesting people use hierarchically organized categorical information when making estimations across scales and across dimensions; and that providing people with more salient category boundary information improves estimation.

In both the event time line and object distance tasks, participants across conditions were relatively accurate in identifying the location of the 1st (Life appears/Pluto) and 4th (Humans appear/Cruithne) events/objects (respectively). This may be because the 1st and 4th events/objects are anchored by the relatively close flanks of the number line itself ('top' and 'bottom'), whereas the 2nd (Dinosaurs appear/Mars) and 3rd (Dinosaurs disappear/Mercury) events/objects (respectively) may not be naturally perceived in these same salient categories; they are located 'somewhere in between'. Consistent with this interpretation, participants from the conventional condition demonstrate more bias in estimation towards the center of the number line than the participants from the hierarchical condition (as seen in the overestimation of the 2nd and 3rd events/objects). This finding is aligned with the three-category representation of geologic time advocated by Trend (2001), as well as predictions of biases towards the middle of these categories by the CAM. However, more research is needed to further identify and characterize categories used in the representation of geologic time and astronomical distances.

Participants across conditions were significantly more accurate on the object distance task (μ error = 33.76mm) than the event time line task (μ error = 45.45mm) ($t(39)=2.85$, $p=.01$). Participants across conditions also were more accurate on the abstract (numeric) magnitude task (μ error = 22.7) compared to the event time line task ($t(39)=5.55$, $p<.001$) and the object distance task ($t(39)=2.96$, $p=.01$). One explanation for this pattern of differences in performance is that temporal, spatial, and abstract magnitudes are represented differently (see Agrillo, Ranpura, & Butterworth, 2010 and Walsh, 2003 for a discussion on a general magnitude system). Alternatively, it may be the case that temporal, spatial, and abstract magnitudes are all represented in a similar way, but preexisting knowledge (and misconceptions) bias the subjective categories people use to make estimations. For example, consistent with participants being better at the object distance task compared to the event time line task, that geologic time is often neglected in the classroom (Dodick, 2007; Trend, 2001) and learning about the solar system is commonplace, it seems likely participants did have more knowledge of the solar system than geologic time. Related, the first three base analogies (tens, hundreds, thousands) may be differentially familiar to participants for temporal and spatial magnitudes. While temporal and spatial scales of magnitude were aligned, participants may be more familiar with traveling tens, hundred, and even thousands of miles; whereas participants could have only personally experienced years at the tens scale (no participants were over one hundred years old). Alternatively, mapping the vertical distances onto a horizontal map, and not having an analogous temporal activity, may have contributed to the observed domain differences. Future research should examine unfamiliar scales, both in content and magnitude. One may use an unfamiliar solar system, which would have a different time-course and different celestial objects.

Findings from the abstract magnitude task are consistent with the segmented number line model of scale representation (Ebersbach, et al., 2008; Landy, Silbert, & Goldin, 2012). The segmented linear model posits separate linear functions for familiar versus unfamiliar magnitudes when estimated magnitude is plotted against actual magnitude. Ebersbach and colleagues (2008) found young children had a fairly accurate linear slope for smaller, familiar numbers, and a separate shallower linear slope for larger, unfamiliar numbers. While there were not enough estimations in the current study to carefully characterize the slope function, participants across conditions had a more accurate linear slope for estimations made on the million scale, and, while still linear, were significantly less accurate on estimation on the billion scale (overestimation). More research is needed examining estimations at large scales for detailed modeling of these slope functions.

That the hierarchical condition transferred to estimations about abstract magnitudes, suggests that people use categorical information when making these types of estimations. While there are some studies that look at the subjective categorization of numbers (Laski & Siegler, 2007; Mix, Huttenlocher, & Levine, 2002; Siegler & Robinson, 1982), there has not been previous work mapping the CAM onto number line estimations and scale representation. While more direct and explicit research is needed, we speculate that the CAM could serve as a unifying model for currently competing theories (e.g., logarithmic-to-linear, power function with anchor points, segmented linear). Category boundaries may serve as distinct anchor points, with adults possessing more precise categories (at the individual numbers level) compared with children. Whereas young children may have many numbers in one “big” or “unfamiliar” category, adults may possess counting strategies for numbers within “unfamiliar” scales. Thus, the CAM offers an account for the overestimation of unfamiliar magnitudes that maintain linearity within the scale. More extensive research is needed to identify types of categories used in scale representation to see if a CAM can predict the changing pattern of bias in number line estimations that occurs with development.

An implication of the current findings is the hierarchical alignment model is an effective way to teach about scales outside of human perception. Understanding scale information is important, as fundamental concepts in many disciplines require understanding of scales outside of human experience. “Size and scale” have been identified by the new *National Research Council Framework for K-12 Science Education* (2011) and the *Benchmarks for Science Literacy* (AAAS, 1993) as a fundamental and unifying theme of science education. Having a linear representation of scale is predictive of performance on a range of standardized tests in mathematics (Siegler & Booth, 2004). Unfortunately, understanding large scales is difficult (e.g., Libarkin, et al., 2005; Tretter, et al., 2006). Undergraduate students, even those in science, technology, engineering, and mathematics majors, have difficulty mastering concepts of size and scale (Drane et al., 2008). While people are fairly accurate on identifying correct sequences, they fail to understand the magnitude between the events (Tretter, et al., 2006) and objects (Jones, et al., 2008). By providing a salient internal structure of magnitude boundaries, the hierarchical alignment activity may be an effective classroom tool to help foster a linear representation of scales like geologic time and astronomical distance.

Acknowledgments

This research was funded in part from National Science Foundation Grant SBE-0541957 and National Science Foundation Grant SBE-1041707 that both support the Spatial Intelligence and Learning Center (SILC).

References

Agrillo, C., Ranpura, A., and Butterworth, B. (2010). Time and numerosity estimation are independent: behavioural evidence for two different systems using a conflict paradigm. *Cognitive Neuroscience, 1*

American Association for the Advancement of Science (AAAS). (1993). *Benchmarks for science literacy*. NewYork: Oxford University Press.

Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: evidence against a representational shift. *Developmental Science, 14*.

Clark, H.H. (1973). Space, time, semantics, and the child. In T.E. Moore (Ed.), Cognitive development and the acquisition of language. New York: Academic Press.

Dehaene, S. & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition: International Journal of Cognitive Science, 43*(1)

Dodick, J. (2007). Understanding evolutionary change within the framework of geological time. *McGill Journal of Education, 42*(2).

Drane, D., Swarat, S., Hersam, M., Light, G., & Mason, T. (2008). An evaluation of the efficacy and transferability of a nanoscience module. *Journal of Nano Education*.

Ebersbach, M., Luwel, K., Frick, A., Onghena, P., & Verschaffel, L. (2008). The relationship between the shape of the mental number line and familiarity with numbers in 5- to 9-year-old children: Evidence for a segmented linear model. *Journal of Experimental Child Psychology, 99*.

Huart, J., Corneille, O., & Becquart, E. (2005). Face-based categorization, context-based categorization, and distortions in the recollections of gender ambiguous faces. *Journal of Experimental Social Psychology, 41*.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review, 98*.

Huttenlocher, J., Hedges, L. V., & Prohaska, V. (1988). Hierarchical organization in ordered domains: Estimating the dates of events. *Psychological Review, 95*.

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General, 129*.

Gentner, D. (2001). Spatial metaphors in temporal reasoning. In M. Gattis (Ed.), Spatial schemas and abstract thought. Cambridge, MA: MIT Press.

Jones, M. G., Tretter, T., Taylor, A., & Oppewal, T. (2008). Experienced and novice teachers' concepts of spatial scale. *International Journal of Science Education, 30*(3).

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67*.

Landy, D., Silbert, N., Goldin, A. (2012). Getting off at the end of the line: the estimation of large numbers. *2012 Annual Meeting of the Cognitive Science Society Conference Proceedings*.

Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development, 76*.

Libarkin, J.C., Anderson, S.W., Dahl, J., Beilfuss, M., & Boone, W. (2005). Qualitative Analysis of College Students' Ideas about the Earth: Interviews and Open-Ended Questionnaires. *J. of Geoscience Edu., 53*(1).

Libarkin, J.C., Kurdziel, J.P. & Anderson, S.W. (2007). College student conceptions of geological time and the disconnect between ordering and scale. J*ournal of Geoscience Education*, 55: 413-422.

Miller & Brewer. (2010). Misconceptions of Astronomical Distances. *International Journal of Science Education, 32*(12).

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). *Quantitative development in infancy and early childhood*. New York: Oxford Press.

National Research Council. (2011). *A Framework for K-12 Science Education.* Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, DBASSE. Washington,DC: The National Academies Press.

Opfer, J. E., Siegler, R. S., & Young, C. J. (2011). The powers of noise-fitting: Reply to Barth and Paladino. *Developmental Science, 14*.

Resnick, I., Shipley, T., Newcombe, N., Massey, C., Wills, T. (2012). Examining the Representation and Understanding of Large Magnitudes Using the Hierarchical Alignment model of Analogical Reasoning, *2012 Annual Meeting of the Cognitive Science Society Conference Proceedings*.

Roberson, D., Damjanovic, L. & Pilling, M. (2007). Categorical Perception of Facial Expressions: Evidence for a 'Category Adjustment' model. *Memory & Cognition, 35*

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*.

Siegler, R. S., & Robinson, M. (1982). *The development of numerical understandings.* In H. W. Reese & L. P. Lipsitt (Eds.), Advances in child development and behavior: Vol. 16. New York: Academic Press.

Thompson, C., & Opfer, J. (2010). How 15 hundred is like 15 cherries: Effect of progressive alignment on representational changes in numerical cognition. *Child Development, 81*(6)

Trend, R.D. 2001, Deep Time Framework: a preliminary study of UK primary teachers' conceptions of geological time and perceptions of geoscience. *Journal of Research in Science Teaching, 38*(2)

Tretter, T. R., Jones, M. G., Andre, T., Negishi, A., & Minogue, J. (2006). Conceptual boundaries and distances: Students' and experts' concepts of the scale of scientific phenomena. *Journal of Research in Science Teaching, 43*(3).

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *TRENDS in Cognitive Sciences, 7*(11)