How the Appearance of an Operator Affects its Formal Precedence

David H. Landy (dlandy@uiuc.edu)
Department of Psychology, 603 E. Daniel Street
University of Illinois Champaign, Illinois USA

Michael N. Jones (jonesmn@indiana.edu)
Department of Psychology, 1101 E. 10th Street
Indiana University, Bloomington, Indiana USA

Robert L. Goldstone (rgoldsto@indiana.edu)
Department of Psychology, 1101 E. 10th Street
Indiana University, Bloomington, Indiana USA

Abstract

Two experiments test predictions of a visual process-driven model of multi-term arithmetic computation. The visual process model predicts that attention should be drawn toward multiplication signs more readily than toward plus signs, and that narrow spaces should draw gaze comparably to multiplication signs. Although both of these predictions are verified by behavioral response measures and eye-tracking, the visual process model cannot account for patterns of early looking. The results suggest that people strategically deploy visual computation strategies.

keywords; cognitive arithmetic; eye-tracking; formal reasoning; mathematical cognition

Introduction

Although mathematics is often described as a universal language, and the concepts involved in mathematics can be expressed in many different ways without altering their meaning, actual mathematical statements are almost always written using one highly specialized notation (Arabic numerals together with symbols like +, x, ÷ and =). Learning this notation is a difficult task (Nathan, Alibali & Koedinger, 2004), and failures to understand the formal notation are not always easy to distinguish from conceptual misunderstandings. Since processing symbols via formal rules is central to cognition, understanding how people learn to use mathematical notation promises to shed light on the fundamental structures that implement human reasoning.

It may appear at first glance that formal notation is simply a symbolic shorthand. For instance, $1+1=2$ may be read *ad verbum* as "one plus one equals two". However, formal notation also involves idiosyncratic rules for interpreting particular kinds of structures. For instance, it is not clear without instruction how to read such non-linear expressions as summations with upper and lower bonds. Even within purely linear examples, formal notation involves idiosyncratic interpretive rules. The English-language phrase "two plus two times five", for example, might mean "twelve" or "twenty", the formal expression $2+2\times5$ refers uniquely to 12, because of an "order of operations" rule which specifies that multiplications are resolved before additions. These rules are not part of the conceptual structure of mathematics, nor are they contained in the meaning of the operations. Instead, they are idiosyncratic parts of the common mathematical symbology, which must be mastered by learners.

One might posit that these processes of intepretation, like the concepts of mathematics, are mentally represented as rules. These rules are applied by a post-perceptual parsing process which turns the linear, string-based representation into a hierarchical mental representation, suitable for computation (Keodinger & MacLaren, 1998). The idea explored here is that, instead, evaluation of mathematical expressions typically follows a very different course: we suggest that mathematical interpretation is accomplished *via* mechanisms of attentional control and gaze direction on literal expressions. On this view, formal notations are far more than a symbolic shorthand—they are physical systems which have become very well-suited to the human perceptual structures, so that what is mathematically correct will be perceptually natural. Seen this way, mathematical reasoning is a form of extended cognition (Clark, 2007), where correct formal behavior results not from a mentally representated rule, but from the interaction of typical bodily engagement and the constructed external environment.

The major focus of this paper will be to explore one particular manifestation of the embodied approach to formal syntax parsing, which we will call the salience model of simple arithmetic, or just the salience model. One computational instantiation of this model appears in Landy (2007), but the idea is relatively general. In broad terms the salience model posits that single-operation arithmetic computation occurs largely automatically when a person attends to a particular sub-problem, at a rate that combines the difficulty of that problem (for instance, divisions and multiplications would tend to take longer than additions of equally sized operands) with the visual regularity of that problem. This latter criterion suggests that, for instance, $8 \;\;\times\;\; 3$ would be a less regular instance of "eight times three" than would 8×3. Similarly, in $9+2 \quad + \quad 6$, the left sub-problem would form a better group than the right because the 9 and 2 are physically closer than the 2 and 6. The salience model also assumes that multiplications form better groups than do additions, and therefore that multiplications should be more salient than additions. Activity proceeds, in the implementation described in Landy (2007), because once a sub-problem is solved, it is visually inhibited, and the result visually imagined in its place.

In the salience model, multiplications should be privileged not just in mathematical or computational contexts, but whenever multiplications and additions are competing for attention. This prediction is tested in Experiment 1. Moreover, multiplication signs and spatial proximity act in the same way, drawing attention and processing power to an item, and facilitating its solution. In the context of simple arithmetic problems over addition and multiplication, this implies that problems in which multiplications are more narrowly spaced than additions should be easier to solve than the reverse, which has previously been reported (Kirshner, 1989; Landy & Goldstone, 2007A); hence, we will refer to such problems as "consistent", as opposed to "neutral" (with uniform spacing) or "inconsistent" (additions more narrowly spaced than multiplication) problems. It also implies that both multiplications and narrowly spaced problems should draw attentional resources more strongly than additions and widely spaced problems. This hypothesis is tested in Experiment 2, which uses gaze position and fixation duration as estimates of attentional resources. Several prior studies of eye-movement in arithmetic have been performed. These studies, however, have tended to look at vertically arranged stimuli with a single operation type (Suppes, 1990), or to use regular stimuli, where looking patterns do not need to vary (Salvucci 1998), and none to our knowledge explores the role of spacing in computation.

The salience model can be usefully contrasted with a default serial model. Any plausible serial model of arithmetic, for instance, should assume that multiplications are evaluated before additions, and therefore that attention (and gaze) will be focused on multiplications early in trials, and on additions later. Such a model makes no predictions about how people will treat narrowly- or widely- spaced sub-problems, however, and furthermore provides little reason to suppose that attention would be drawn more rapidly to multiplications than to additions prior to encoding the problem. In the salience model, by contrast, attention should immediately tend toward narrow spaces and multiplications. The default serial model is not attributable to any particular researcher, nor do we mean to suggest that it is generally tacitly assumed. The intended role of the default serial model is not as a straw man but as a baseline.

Experiment 1: Attentional Features of •, ×

Method

Forty-eight Indiana University undergraduates received partial course credit for participation. Three participants were removed for failing to reach a criterion of 80% overall accuracy. Including these participants did not affect the significance or pattern of results.

All stimuli consisted of valid mathematical expressions with one addition and one multiplication operation. The three operands were single digits between 2 and 9, and were printed in the LeHei Pro font on Macintosh computers. The plus sign was a simple vertically and horizontally symmetric +. For 19 participants in the "cross" condition, the multiplication sign was a ×; the remaining 26 participants in the "dot" condition saw a •used for multiplication. The cross sign was identical to the plus sign, but was rotated 45° All of the operand symbols were 14mm across; operands were separated by 50mm.

Participants were instructed to press a button corresponding to the side of the expression on which a target sign appeared. The target sign alternated between the addition sign and the multiplication sign in blocks of 20 trials.

Results

Accurate-trial response times are summarized in Table 1. Participants responded more quickly when identifying multiplication signs than addition signs ($F(1,43)=12.4$, $p<=.01$); this difference was also significant within each symbol condition. Further, there was also a main effect of symbol condition: participants responded more quickly in the dot than the cross condition ($F(1,43)=85$, $p<.001$). There was no significant interaction between task and symbol condition ($F(1,43)=1.45$, $p\sim.23$).

Table 1: Response time (and accuracy) in Experiment 1

Condition	Task	
	Addition	Multiplication
Cross	947±19ms (.95±.01)	916±18ms (.98±.01)
Dot	714±17ms (.93±.03)	698±19ms (.97±.02)

Discussion

As predicted by the salience model, people tend to respond more quickly when searching for multiplications than when searching for additions. Note, however, that the salience model does not provide an account of why multiplication searches should be more rapid than addition searches. It is possible that this is the result of long-term training on mathematical tasks. It is also possible that the visual features of the multiplication signs make them stand out better from numbers than do those of the addition sign.

Although Experiment 1 demonstrated a response bias in favor of the multiplication on a non-mathematical task, it does not demonstrate that multiplications attract attention more readily. It might be that both problem types are equally attractive, but that responding to additions is generally inhibited. Experiment 2 provides a convergent measure of the attractiveness of multiplications, by measuring eye-position during arithmetic problem-solving.

Experiment 2: Eye-Tracking

Method

Participants were 13 undergraduate students from Indiana University, who received partial course credit for participation. The experiment lasted about 50 minutes.

Participants were shown a set of 144 simple two-operator arithmetic problems, and asked to compute their value and say the result out loud as they pressed a button signaling completion. Eye movements were recorded until participants pressed a button indicating completion. Participants were reminded of the order of operations before beginning the experiment.

Operands ranged from in magnitude from two to nine; the operations were addition or multiplication. Thus, problems could have the operator structure *plus-plus*, *times-times*, *times-plus*, or *plus-times*. Each was shown equally often.

Each problem was displayed twice, once each in two of four possible spacing conditions: In the *narrow-wide* or NW condition, the left operation was more narrowly spaced than the right, as in 2*3 + 4; in the *wide-narrow* or WN condition, the reverse was true, as in 2 * 3+4. In the *wide-wide* or WW condition, both terms were widely spaced, and finally in the *narrow-narrow* or NN condition, both operators were narrowly spaced. Problems were presented equally often in each spacing condition and operator group.

Each operand and operator was 16mm wide. Narrowly spaced operands were placed 31mm apart; widely spaced operands were placed 101mm apart. In both cases, operators were placed equidistant from each operator. Eye positions were recorded with an Eyelink 1000/2K Desktop Mount System from SR Research running at a temporal resolution of 250Hz. We used monocular tracking locked to the participant's right eye only. Each stimulus display subtended a maximal visual angle of 23 degrees.

Measuring Gaze Position In considering gaze position, we are more concerned with the horizontal than the vertical position. However, there are two plausible ways to measure horizontal position. We might measure the distance from the expression center, that is, the midpoint between the two outer operands. Alternatively, we might measure the displacement of gaze to the right or left of the central term (see Figure 1). When expressions are uniformly spaced, these two measures coincide, but when the expression is either in the narrow-wide or wide-narrow condition, the measures differ. Neither measure is perfect: for this report, we shall measure displacement from the central term when the mathematical algorithm is being investigated, and use deviation from the expression midpoint when marking the effect upon gaze of some dependent manipulation. This is a reasonable choice because displacement from the central term determines which sub-problem is fixated, but the expression midpoint constitutes a natural zero-point for overall gaze-direction. Most reported contrasts are unaffected by using the alternate measure, and all patterns are qualitatively similar.

In this study, we used three measures of eye-movements: total gaze time during a portion of the trial, mean fixation duration, and horizontal position of gaze. The first two of these provide a measure of processing. Longer fixations have been associated with more difficult or more conceptual material, while time allocated to an area may generally correspond to a center of attention.

Because the most distinctive predictions of the salience model involve initial looking patterns, we performed a separate analysis of the first fixation after the pre-trial cue.

Results

Overall Performance Participants were quite accurate on these simple problems; mean accuracy over all trials was 93.8%. Participants timed out on 2.2% of trials. These trials were removed from all analyses. After removing these trials, median RT was 2,890ms. The median number of fixations per problem was 9.

First looks tended to be directed slightly left of the center of the expression (mean displacement from expression center: -13.4 ± 3mm, where the negative denotes that the mean was left of center), which on uniformly spaced trials was nearest the middle operand.

Effect of Consistency on Performance The small number of subjects limits the conclusions that can be drawn about the relationship between consistency and accuracy and response time, but the trends are substantially compatible with those reported in Landy and Goldstone (2007A). Items tended to be solved both more accurately and more quickly in the more consistent format (2701ms vs 2838ms), but neither difference was significant (Accuracy: Mean = 94.7% vs. 89.7%; $F(1,12)=1.32$, $p\sim.27$; Response Time: Mean = 2701ms vs. 2838ms; $F(1,12)=4.0$, $p\sim.07$).

Overall looking patterns Figure 2 shows overall looking across the fraction of the trial completed for *times-plus* and *plus-times* trials (*plus-plus* and *times-times* trials look very similar to *times-plus* trials). Figure 2 is quite compatible with the baseline expectation that people tend to solve sub-problems serially: they do the multiplication in the first half

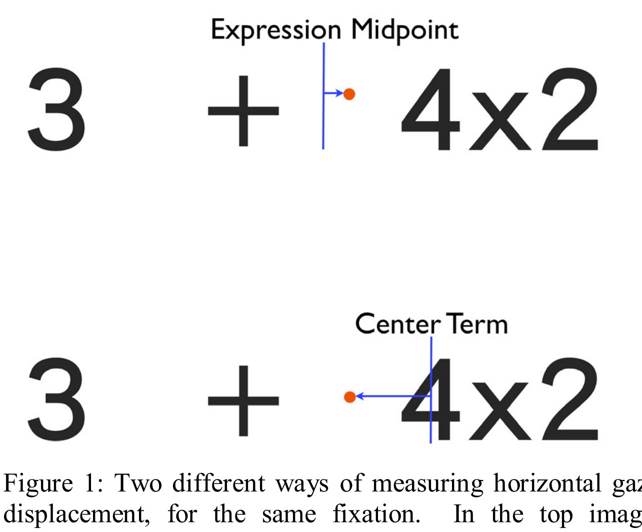

Figure 1: Two different ways of measuring horizontal gaze displacement, for the same fixation. In the top image, displacement is measured from the spatial midpoint of the problem. In the bottom, it is measured by displacement from the structural middle. The sign of the latter measure corresponds to the sub-problem being fixated.

Table 2: Mean Fixation (Total Gaze) Duration (ms) as a function of trial and operator of target sub-problem

Portion of trial	Sub-problem Addition	Multiplication
First half	217±13 (415±33)	274±22 (694±65)
Second Half	316±20 (697±50)	312±29 (560±61)

Table 3: Mean Fixation (Total Gaze) Duration (ms) as a function of portion of trial and width of target sub-problem

Portion of trial	Sub-problem spacing Wide	Narrow
First half	207±15 (405±31)	280±29 744±47
Second Half	284±17 (574±51)	327±19 727±42

of the trial, and the addition in the second half. For this reason, analyses of overall gaze duration split the trial into two halves, and sum the gaze time over each half.

Separate analyses of variance (ANOVAs) of trials in which the sub-problems differed (*times-plus* and *plus-times* problems) were performed to evaluate total gaze toward multiplications and the duration of individual fixations at multiplications. In the former, trial half and fixation target served as dependent measures; in the latter, time of fixation was used as a continuous measure. Because the results were identical, the descriptions are collapsed.

These analyses revealed that while there was an expected interaction between sub-problem fixated and gaze (see Table 2; F(1,12)=19.5, p<.001 for total gaze, F(1,12)=8.5, p<.05 for individual fixations.), there was no significant difference by either measure between total duration of gaze at additions and multiplications (Gaze: F(1,12)=1.18, p~.3; Fixation: F(1,12)=1.4, p~.27).

A separate pair of analyses was performed on all fixations that differed in spacing, to explore the relationship between spacing and total looking time over the course of the trial. Again, time of look was binned over trial half for total gaze. As predicted by the salience model, the analysis revealed that participants spent more time overall gazing at narrowly-spaced sub-problems than at widely spaced, and that individual fixations were longer (Gaze: F(1,12)=14.9, p<0.01; Fixations: F(1,12)=30.8, p<.001). Also, there was a

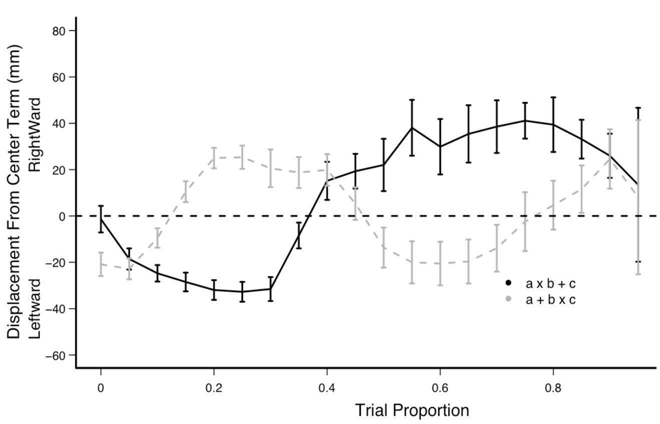

Figure 2: Mean gaze position (positive is rightward) for *times-plus* and *plus-times* problems.

significant interaction between width and trial half (Gaze: F(1,12)=10.0, p<.01; Fixation: F(1,12)=15.0, p<.01); early in the trial participants fixated more strongly on narrowly spaced elements.

First fixations. To test the prediction that eye-gaze will be systematically biased toward the multiplication sign, the mean deviation of gaze from the expression midpoint was computed for each participant. These means were not significantly different from 0 overall (t(12)=0.14, p~.9); however, there was a significant effect of spacing uniformity on multiplication-sign looking. When spacing was uniform, initial fixations were biased in the direction of the multiplication sign (t(12)=2.9, p<.05; see Table 4), but when spacing was non-uniform (i.e., when one sub-problem was narrowly spaced and the other widely spaced), looks were marginally biased in the direction of additions (t(12)=1.99, p~.07; the difference was also significant, t(12)=3.1, p=.01). Early looks toward additions tended to be shorter than looks that fell inside a multiplication (M=28ms, t(12)=2.6, p<.05).

A complementary analysis was performed to evaluate the impact of narrow spacing on gaze direction, with analogous results. Overall gaze direction was not biased toward or away from narrow spaces (t(12)=1.2, p~.25), but within problems whose operators were the same (*plus-plus* and *times-times* trials), bias was significantly in the direction of narrow spaces (t(12)=4.4, p<.001, see Table 5), and within trials with different operator structures, gaze was systematically directed toward wide spaces (t(12)=3, p=.01). Early looks toward wide spaces were shorter than looks that fell inside narrow spaces, (M=54ms, t(12)=6.0, p<.001).

Eye-movements and Consistency Consistency could be thought of as an interaction between the effects discussed above, but it is simpler for expository purposes to code consistency separately, and analyze its effect as an independent property (the full 4-way analyses are qualitatively similar in all cases). Figure 6 displays the mean gaze position over the entire trial for consistent vs. inconsistent displays. In this graph, in addition to a general tendency in both conditions to move toward the addition over the course of the trial, there is an evident difference in first glances: first glances in consistent trials fall inside the addition, while those in inconsistent trials fall inside the multiplication. On inconsistent trials, on the other hand,

Table 4: Mean position of first looks, for trials in which the spacing was not uniform.

Operator Structure	Spacing Structure Narrow-Wide	Wide-Narrow
Same	-14.9±6.6mm	-5.9±7.1mm
Different	-6.4±7.6mm	-14.2±6.7mm

Table 5: Mean position of first looks, for trials in which the operators differed.

Spacing Structure	Operator Structure Times-Plus	Plus-Times
Neutral	-20.3±6.3mm	-6.2±8.7mm
Non-Neutral	-13.3±8.4mm	-14.5±7.4mm

gaze initially fell within the multiplication, but moved quickly over to the addition, and then part-way back, as though participants were glancing toward the addition before computing the multiplication. Along the same vein, notice that gaze is more strongly patterned in the consistent condition: on inconsistent trials, gaze tends to be centrally located throughout the trial. (This pattern was confirmed by three-way ANOVAs of gaze duration and fixation duration over trial time, consistency, and target operator. Total gaze: $F(1,12)=9.1$, $p=.01$; Fixations: $F(1,12=6.6$, $p<.05)$.

Discussion

Experiment 2 revealed significant modulation of behavior based on overall expression structure and physical spacing. People treat spaces and multiplications very similarly, as predicted by the salience model. Further, people tend to focus on narrow spaces (and multiplications) early in a trial, and move toward wider spaces (and additions) toward the end. Participants spent more time looking at narrowly-spaced sub-problems, and individual fixations directed toward them were longer, confirming the hypothesis that narrowly-spaced problems draw attention (although this result was not replicated for multiplications). Overall, Experiment 2 confirmed the prediction that narrow spacing and multiplications have similar effects on attention.

The pattern of results in first fixations is more complicated, and less compatible with the predictions of the salience model. The model predicts that gaze should be

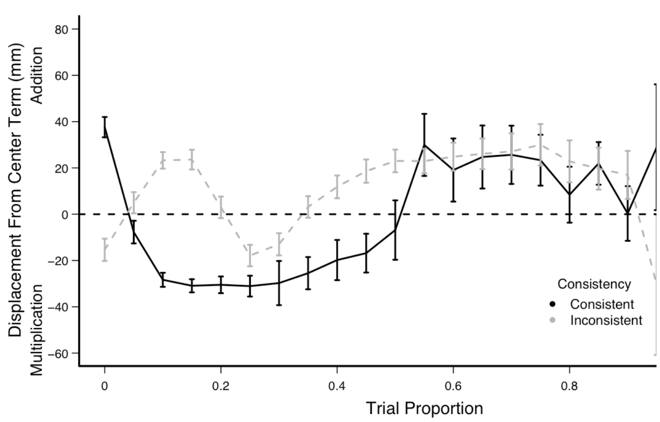

Figure 6: Mean fixation position over trial, for consistent and inconsistent stimuli. Here negative denotes the high precedence side (i.e., the side with the multiplication on it).

directed first toward multiplications and narrow spaces, but this turned out to be true only for stimuli in which spacing was uniform, and the operations were the same, respectively. The hypothesis that narrowness and multiplication are processed similarly provides a simplifying interpretation of these data. In this interpretation, there are two kinds of expressions (call them “homogeneous” and “non-homogeneous”), and they have preferred computation strategies. Homogenous expressions tend to be equally spaced, and to have identical operators. In these equations, people tend to focus immediately on high-salience elements: multiplications and narrow spaces. In non-homogeneous expressions—ones with different operators and different spacings—people tend to focus briefly on the lower-priority items—wide spaces and additions—before turning to the additions. Spatial consistency, in this interpretation, is helpful not (just) because it helps people focus on the multiplications throughout the first part of the trial, but largely because it helps people focus quickly on the addition signs.

This interpretation also makes the following prediction: despite the fact that in our formal notation, multiplications are more salient than additions, people should perform better in a system in which the high-precedence operations were less salient than the low-precedence. If so, then it seems that, in contrast to the predictions of the salience model, people select strategies driven by the overall structure of the expression, rather than following default biases to attend to narrow spacing and high-precedence. Further work will be needed to evaluate, verify, and extend the pattern of first fixations found here.

General Discussion

The results presented here were intended as a test of a visually-driven model of computation. Most basically, the model predicted that multiplication signs would attract response-attention over addition signs on non-mathematical judgments, and this prediction was supported. Furthermore, the model predicted that when evaluating arithmetic expressions people would treat narrowly spaced problems like multiplications, and this was largely supported across a variety of measures. However, the model also predicted that people would initially look toward both narrow spaces and multiplications; this prediction was not supported. In contrast to the predictions of Landy (2007), the pattern of responses is more consistent with the existence of a distinct pre-computation stage, in which the low-precedence operators must sometimes be examined directly.

One of the values of a computational model is that it makes specific, falsifiable predictions. The visually-driven model of arithmetic computation has done so. The results presented here seem to contradict the model of arithmetic processing which was their initial motivation, but they do not imply that visual properties are unimportant to reasoning with formal expressions. On the contrary, both eye-tracking and behavioral measures reinforce the conclusion that visual properties of formal expressions play a central role in their interpretation, supporting the view that high-level formal cognition is closely tied to physical and spatial reasoning (Anderson et al, 2007; Landy & Goldstone, 2007; Dorfler, 2002).

Clearly, further work is needed to evaluate the relationship between operator symbols and equation interpretation. The predictions of the interpretation presented here are difficult to test with simple two-operator expressions. Future work will include measuring gaze-position on more complex compound expressions. The complex relationship between salience and computation algorithms has implications for both education and interface-design.

Understanding the fine structure of the processes people use to parse novel arithmetic equations is valuable because it provides constraints on models of formal reasoning. Endress et al (2005; see also Pothos, 2006) argued that human learning of formal structures is not general, but must instead bootstrap itself in the perceptual processes that act upon representations of those structures. The current work supports this perspective, by showing that even on highly studied formal tasks, people are affected by visual structure. At best, the current study illustrates how rule-guided parsing shares processes with spatial object perception online during problem solving. At the least, this research suggests that the mapping between operator salience, spacing, and order of operations, is not as straightforward as it may have appeared at first glance.

Acknowledgments

This research was funded by Department of Education IES grant R305H050116, and by AFOSR Grant # FA9550-07-1-0147. The authors would like to thank Jason Dawson and Thomas Hills for valuable comments and feedback.

References

Anderson, J. R., Qin, Y., Jung, K. J., & Carter, C. S. (2007). Information-processing modules and their relative modality specificity. *Cognitive Psychology, 54,* 185-217.

Clark, A. 2006: *Natural born cyborgs*. New York: Oxford University Press.

Dorfler, W. (2002). Formation of mathematical objects as decision making. *Mathematical Thinking and Learning, 4,* 337-350.

Endress, A. D., Scholl, B. J., & Mehler, J. (2005). The role of salience in the extraction of algebraic rules. *Journal of Experimental Psychology: General, 134,* 406-419.

Kirshner, D. (1989). The visual syntax of algebra. *Journal for Research in Mathematics Education, 20,* 274-287.

Kirshner, D. & Awtry, T. (2004). Visual salience of algebraic transformations. *Journal for Research in Mathematics Education, 35,* 224-257.

Koedinger, K. R., & MacLaren, B. A. (2002). Developing a pedagogical domain theory of early algebra problem solving. *CMU-HCII Tech Report 02-100*.

Koedinger, K. R. & Nathan, M. J. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *Journal of the Learning Sciences, 13,* 129-164.

Landy, D. (2007). *Formal Notations as Diagrams of Abstract Structure.* Doctoral dissertation, Indiana University Bloomington, August 31, 2007.

Landy D. & Golstone R. L. (2007A). The alignment of ordering and space in arithmetic computation. *Proceedings of the 29th Annual Meeting of the Cognitive Science Society*.

Landy, D. & Goldstone, R. L. (2007B). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition. 33*(4) 720-733 .

Landy, D. & Goldstone, R. L. (2008). Formal notations are diagrams: Evidence from a production task. *Memory and Cognition, 35*(8) 2033-2040.

Pothos, E. M., Chater, N., & Ziori, E. (2006). Does stimulus appearance affect learning? *American Journal of Psychology 119* (2) pp. 277-301.

Suppes, P. (1990). Eye-movement models for arithmetic and reading performance. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes*, 455-477. Amsterdam: Elsevier.

Salvucci, D. D. (1998). Interpreting eye movements with process models. *CHI 98 Conference Summary on Human Factors in Computing Systems*, 66-67.