Braithwaite, D. W. (2019). *Challenges of Modeling Continuity and Change in Children's Seriation and Ordinal Understanding*. [Peer commentary on the article "The Development of Size Sequencing Skills: An Empirical and Computational Analysis" by M. McGonigle-Chalmers and I. Kusel]. *Monograph Matters*. Retrieved from https://monographmatters.srcd.org/2019/11/12/commentary-braithwaite-84-4/

Challenges of Modeling Continuity and Change in Children's Seriation and Ordinal Understanding

David W. Braithwaite, Ph.D.
The Florida State University
braithwaite@psy.fsu.edu

The monograph by McGonigle-Chalmers and Kusel (2019) describes a series of experiments providing evidence for discontinuity between ages 5 and 7 in children's performance on seriation and ordinal understanding tasks, and offers a theoretical account of this development. Central to their argument is the proposal that discovery of advanced strategies results from creation of new representational structures that are made possible by developmental increases in working memory (WM). In my commentary, I evaluate two aspects of the authors' proposal: (1) the mechanisms underlying children's discovery of new strategies, and (2) the role that working memory plays in driving strategy discovery. I focus on the computational models that are presented in the monograph as implementations of its central theoretical proposal. I conclude by discussing general implications regarding continuity and discontinuity in development.

Mechanisms of Strategy Discovery

McGonigle-Chalmers and Kusel argue that a qualitative change occurs in how children perform size seriation tasks between the ages of 5 and 7. The question of how this happens is among the most important questions addressed in the monograph, as it relates to a central issue in cognitive development: how children discover and adopt more advanced strategies. The computational models described in Chapter III are presented as an answer to this question. The most advanced seriation model, the principled search model, is described as having "*discovered* [emphasis added] that a heuristic search across many possible actions is less efficient than an algorithm that selects stimuli based on the principled iteration of a 'select smallest difference' rule" (p. 80).

However, contrary to this claim and similar claims throughout the monograph, the principled search model does not discover anything—nor do any of the other models. Instead, each model employs strategies that were hard coded by the model designers. Discoveries occur in the space between models, such as in the shifts from the heuristic search model to the transitional model and from the transitional model to the principled search model. The monograph, however, does not specify the mechanisms that underlie these discoveries. This omission considerably weakens the models' utility for advancing understanding of strategy discovery.

In principle, computational cognitive models can simulate strategy discovery. One approach that has been employed towards this end is to endow a model with (1) a set of primitive operators representing simple cognitive processes; (2) a method of combining operators to form representations of more complex cognitive processes; and (3) a method of evaluating and selecting among these combinations of operators according to how well they achieve some goal. These features can enable models to develop capabilities beyond those initially granted to them without the model designers explicitly specifying those capabilities.

In an early example, Shrager and Siegler (1998) employed the approach just outlined to simulate children's discovery of strategies for adding two small whole numbers. In their model, SCADS (Strategy Choice And Discovery Simulation), strategies are represented as sequences of simple operators, such as COUNT ALL FINGERS. The model develops successively more advanced strategies by creating new sequences of operators that calculate sums more efficiently than previous strategies. New sequences of operators are generated by heuristics, such as a heuristic to eliminate redundant steps within a sequence, and are tested against metacognitive constraints, such as the constraint that both addends must be represented. Simulations demonstrated that SCADS discovered the same strategies as children in the same order as children; also like children, SCADS did not adopt illegal strategies such as adding an addend to itself. An extension of SCADS based on similar assumptions simulated children's discovery of a shortcut strategy for solving three-term arithmetic problems (Siegler & Araya, 2005); similar approaches have also been employed to simulate strategy discovery using neural networks (Anumolu, Bray, & Reilly, 1997).

In a more recent example, Piantadosi, Tenenbaum, and Goodman (2012) proposed a Bayesian model of children's acquisition of numerical concepts. In the model, numerical concepts are represented in the model as functions, expressed in a formal language (lambda calculus), that map sets to number words. The model begins with several primitive operators and initially uses these to construct functions that are analogous to the limited numerical knowledge of very young children, like "if the set is a singleton, then return the word 'one'; otherwise, return nothing." After sufficient training, the model constructs functions that are analogous to the more sophisticated knowledge of adults, like this recursive function: "if the set is a singleton, then return the word 'one'; otherwise, return the number word that comes after the word you would return for a set with one fewer element." Piantadosi et al. (2012) described their approach to modeling discovery as *compositionality*: "Learning may create representations from pieces that the learner has always possessed, but the cognitive pieces may interact in wholly novel ways." At a high level, this approach strongly resembles that of Siegler and colleagues. Composition has been proposed as a fundamental cognitive mechanism not only for strategy discovery, but also for skill learning in general (Anderson, 1982).

A similar compositional approach might help to explain the transition from trial and error seriation to principled size seriation in the present context. Fleshing out this possibility would require answering several questions: What operators are already available to children who perform seriation tasks by trial and error? How are these operators selected and combined to form new strategies? In particular, how could the principled search strategy be formed by combining simple, cognitively plausible operators? What constraints are used to evaluate new strategies? What is the mechanism by which new strategies eventually replace old ones? Implementing answers to these questions in a formal model, and demonstrating that the model can generate behavioral patterns similar to those of children at different ages, would strengthen the theoretical contribution of the present monograph.

The Role of Working Memory

The view that increases in WM capacity enable conceptual development was central to the theory of central conceptual structures (Case & Okamoto, 1996), and this aspect of the theory has recently received empirical support (Morra, Bisagno, Caviola, Delfante, & Mammarella, 2019; see also van der Ven, Boom, Kroesbergen, & Leseman, 2012). The novel contribution of the present proposal is to describe precisely a mechanism by which increases in WM could facilitate advances in seriation and ordinal understanding—that is, by enabling the creation of representational structures, called “slots,” that are required by more advanced strategies. The transitional model in Chapter III formalizes this mechanism: WM is represented by a parameter called *WMAvailability*, and larger values of this parameter increase the likelihood of the model creating slots while performing the seriation task.

However, a weakness of the proposal as it is developed in the monograph is that it does not describe the role of WM in performing seriation tasks in the absence of the aforementioned representational and strategic changes. Seriation under these conditions is described in the monograph by the heuristic search model. This model is unaffected by the value of the *WMAvailability* parameter. Furthermore, the model does not include any other parameter or mechanism reflecting variation among individuals in WM capacity. Therefore, the model in principle cannot predict effects of WM capacity on size seriation among children who employ the heuristic search strategy. Most surprisingly, the heuristic search model also cannot predict effects of WM capacity on color sequencing performance, even though the color sequencing task is described in the monograph (p. 52) as “an independent assessment of working memory in a serial learning context,” and the heuristic search model is the only model that is used in the monograph to simulate color sequencing.

Addressing these concerns would require creating a version of the heuristic search model that explicitly incorporates effects of WM capacity. I will refer to this hypothetical model as the “heuristic search model with WM.” This model could have several advantages over the current version. In particular, it likely could explain improvements in color sequencing from age 5 to age 7, as observed in Experiments 1 and 2, as a consequence of increases in WM capacity over this period. It likely also could explain the positive correlation (found among 5-year-olds in Experiment 2) between speed of learning on the color sequencing task and speed of learning on the size seriation task, because higher WM capacity would lead to faster learning on both tasks.

Most importantly, the heuristic search model with WM would presumably predict that increases in WM capacity would cause improvements in size seriation performance even without the representational and strategic changes that are posited in the monograph and described by the transitional and principled search models. Consequently, the heuristic search model with WM would represent a kind of “null hypothesis” against which the models proposed in the monograph could be compared. The question would then become whether adding the transitional and principled search models results in more accurate predictions about the observed age-related changes in task performance than the predictions generated by only the heuristic search model with WM. A considerable increase in descriptive accuracy would be necessary to justify the much greater complexity inherent in using three models, rather than a single model, to simulate children's development. Without evidence for such an increase in descriptive accuracy, it would be premature to reject the null hypothesis: that direct effects of WM increases, without representational or strategic changes, are sufficient to explain the observed changes in behavior.

Implications Regarding Continuity and Discontinuity in Development

The limitations discussed above highlight general considerations for theorizing about and modeling apparent discontinuities in development. First, an apparent discontinuity is not fully explained by positing an unobserved intermediate state of development. The present monograph proposes the transitional model to describe a hypothesized intermediate state between the states described by the heuristic search model and principled search model. However, this proposal begs the question of how children move from the heuristic search model to the transitional model and from the transitional model to the principled search model. Describing a trajectory of discrete states is a valuable first step in modeling development, but a complete account requires descriptions of the mechanisms that underlie transitions between states.

Second, theories that posit discontinuities in the cognitive resources, such as representational structures or strategies, that are used to perform a task should demonstrate their superior ability to explain the empirical phenomena in comparison to plausible alternatives that do not posit such discontinuities. McGonigle-Chalmers and Kusel (2019) provide empirical evidence for an apparent behavioral discontinuity and rightly highlight the need for a theoretical explanation of it. The proposal that this discontinuity results from representational and strategic changes that are driven by increases in WM capacity is plausible. However, the monograph does not demonstrate that this indirect mechanism explains the empirical data better than would a simpler account involving direct effects of WM capacity on performance. Continuous increases in WM and other domain-general cognitive parameters certainly occur and certainly affect performance on a wide range of cognitive tasks. The possibility that these continuous changes alone can account for any behavioral changes observed constitutes a reasonable "null hypothesis" against which hypotheses of developmental discontinuities can and should be tested.

References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*(4), 369–406. https://doi.org/10.1037/0033-295X.89.4.369

Anumolu, V., Bray, N. W., & Reilly, K. D. (1997). Neural network models of strategy development in children. *Neural Networks, 10*(1), 7–24. https://doi.org/10.1016/S0893-6080(96)00078-0

Case, R., & Okamoto, Y. (1996). The role of central conceptual structures in the development of children's thought. *Monographs of the Society for Research in Child Development, 61*(1–2), i+iii-vi+1-295. https://doi.org/10.2307/1166077

McGonigle-Chalmers, M., & Kusel, I. (2019). The development of size sequencing skills: An empirical and computational analysis. *Monographs of the Society for Research in Child Development, 84*(4). https://doi.org/10.1111/mono.12411

Morra, S., Bisagno, E., Caviola, S., Delfante, C., & Mammarella, I. C. (2019). Working memory capacity and the development of quantitative central conceptual structures. *Cognition and Instruction, 0*(0), 1–29. https://doi.org/10.1080/07370008.2019.1636797

Piantadosi, S. T., Tenenbaum, J. B., & Goodman, N. D. (2012). Bootstrapping in a language of thought: A formal model of numerical concept learning. *Cognition, 123*(2), 199–217. https://doi.org/10.1016/j.cognition.2011.11.005

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science, 9*(5), 405–410. https://doi.org/10.1111/1467-9280.00076

Siegler, R. S., & Araya, R. (2005). A computational model of conscious and unconscious strategy discovery. *Advances in Child Development and Behavior, 33*, 1–42. https://doi.org/10.1016/S0065-2407(05)80003-5

van der Ven, S. H. G., Boom, J., Kroesbergen, E. H., & Leseman, P. P. M. (2012). Microgenetic patterns of children's multiplication learning: Confirming the overlapping waves model by latent growth modeling. *Journal of Experimental Child Psychology, 113*(1), 1–19. https://doi.org/10.1016/J.JECP.2012.02.001