**Background**

(Why we need to study children’s problem solving in arithmetic to answer the brain network questions: a. it is a fairly complex but tractable behavior; b. it provides unique developmental perspective)

Cognitive neuroscience is entering an era in which the aim is to understand how complex behavior arises from the dynamic information exchange and control process across the network of brain regions, and how these change with learning and development. Children’s arithmetic, more specifically children's small number addition (hereafter simply called "addition"), provides a unique opportunity to investigate these questions, as the phenomena in this domain are complex and foundational, yet still tractable. Unlike many domains, addition involves a range of cognitive (and often physical) action (Geary et al., 1992; Siegler, 1996), yet it has well-defined structur with formally correct answers, so that the space of possible correct (or plausibly incorrect) algorithms can be clearly mapped out (van Lehn's thesis). Moreover, how children carry out addition, what mistakes they make, and how they transition from pre-addition (usually only knowing counting sequences) through adult-type "pure retrieval" (Ascraft, 1982; Geary et al., 2007; Groen & Parkman, 1972; Siegler & Shrager, 1984; Siegler, 1986), occurs in a relatively narrow cognitive- and neuro-developmental time-window (Svenson & Sjoberg, 1983; Siegler & Shrager, 1984).

(Why we need the novel control/connectionist framework to understand the neurocognitive networks in brain in children? A. solely on neuroimaging data is not informative/inadequate to tell the mechanism; B. previous modeling work generally disconnects with brain data)

As a result of its great educational importance (Geary et al., 1992; Pazza et al., 2010; Price et al., 2007), children's arithmetic has been well-studied via psychological experiments, cognitive neuroscience, computational modeling. On the cognitive neuroscience side, several core brain regions have been delineated as contributing to math competence and processing (Ansari, 2008; Dehaene, et al., 2003; Menon, 2014), including ventral visual stream (e.g., posterior fusiform gyrus; pFG) for decoding number forms, parietal circuits (majorly around inferior parietal sulcus; IPS) for anchoring the visual numerical representations, prefrontal-parietal cortices for manipulating quantity representations in working memory, and medial temporal lobe (MTL), and especially hippocampus for associative memory processing only in children (Cho et al., 2012; Qin et al., 2014; Supekar et al., 2013). Difficulties in math processing in some children (e.g., developmental dyscalculia; DD) are associated with abnormality, either hypo-activation in these brain regions (Ashkenazi et al., 2012) or increased functional connectivity among these regions (Rosenberg-Lee et al., 2014). Studies further suggest that that MTL, left prefrontal and bilateral posterior parietal cortices may relate to uses of different strategies (Cho et al., 2011), and hippocampus seems crucial for the transition from overt to more implicit strategies (Qin et al., 2014). Even given all this detailed knowledge regarding children's arithmetic, we still do not understand how the whole brain system -- the dynamic network of widely-distributed brain regions -- support the organized execution of arithmetic reasoning, nor how these systems, and especially their interactions, change through learning and development.

Computational studies, which have been undertaken for decades (Siegler & Shrager, 1984; Siegler & Shipley, 1995; Shrager & Siegler, 1998 [[Add an Anderson reference!]]), have generally been rendered in purely symbolic, purely connectionist, and in hybrid symbolic/connectionist paradigms. However, due to unavailability of cognitive neuroscientific data at that time, these models are not rich enough to adequately constrain our theorizing about how the brain develops arithmetic skill.

Recently, my colleagues and I have demonstrated that computational neural networks can account for neural activities in human and animal studies, as well as provide mechanistic explanations for dynamic changes and evolution of brain network functions (Chen & Rogers, 2015; Plaut & Buhrmann, 2011; Stoianov & Zorzi, 2012). The primary goal of the proposed research is build from these results and employ a novel systems-control/connectionist framework to understand the interactive dynamics and evolution of arithmetic skill and number sense y modeling the neurocognitive networks in human brain, and thereby bridge the rich infrastructure of cognitive theory, data, and computational experimentation with recent findings from systems neuroscience. Specifically, I propose to build a new computational model, which is a hybrid of classical connectionist models and classical control system models of children's arithmetic and its development, focusing especially on strategy use.

I will be guided in this effort by several leading thinkers with a wide range of expertise very relevant to this specific domain: Vinod Menon is one of the leading systems neuroscientists, and is among the only ones specifically capturing data and theorizing about the development of the brain as a multi-scale dynamical system, especially in math cognition; Jay McClelland is one of the world's leading connectionist modelers and cognitive scientists, and runs a lab at Stanford focused specifically on arithmetic learning; and Jeff Shrager, consulting professor in the Symbolic Systems program here at Stanford, is one of the founders (with Bob Siegler, of CMU) of the field of the computational modeling of arithmetic development. Dr. Shrager wrote three prior, highly cited, computer models of children's arithmetic development, and is also a leading expert on children's and adult's complex cognitive behavior.

**Goal and Hypothesis**

**Goal 1 & 2**: **Develop a neuro-computational model for both behavioral and neurobiological outcomes of arithmetic development in children.** Using connectionist neural network models (NN models), I will establish a simple model to demonstrate face validity of the neuro-computational approach to account for behavioral and neurobiological findings on children’s development of arithmetic skills. ***Hypothesis 1***: The development of arithmetic problem in the NN models solving is accompanied by transition of strategy use, and results in behavioral changes (e.g., in accuracy or response latency). ***Hypothesis 2***: The development of arithmetic problem in the NN models leads to different profiles of changes in neural activities associated with different problem solving strategies.

**Goal 3: Establish the control/connectionist hybrid model to explain arithmetic development through learning. *Hypothesis 3A***: The developmental transition of strategy uses requires a joint collaboration among perceptual, attentional, working memory, associate memory brain systems, and the less efficient strategies requires sequential processing through multiple brain systems but repetition of using these strategies facilitates the use of retrieval strategies. ***Hypothesis 3B***: The order of developing strategy uses through learning is critical: learning the overt and less efficient strategies provides the cognitive and neural basis for retrieval strategy.

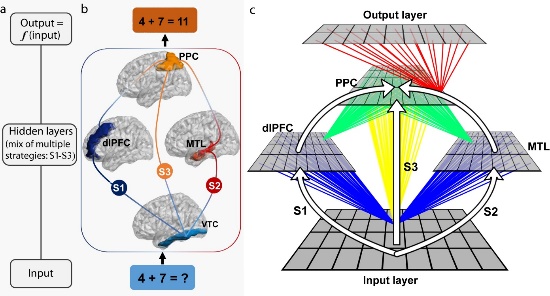
**Goal 4: Explore sources of individual differences in neural basis to explain typical and atypical development**. Based on the extended models from Goal 3, we plan to explain possible neural mechanisms of math difficulties in children with developmental dyscalculia (DD). ***Hypothesis***: Insufficient activation in multiple brain systems, including attentional and working memory can lead to difficulties in establish stable representations for solving arithmetic problems.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (Need further work from here)**

**Experimental methods**

**Aim 1**: **Develop a neuro-computational model to account for behavioral outcomes of strategy transitions throughout development.**

***Neural network model rationale and specification***. The model architecture is depicted in Figure 3. Based on three distinct neurocognitive pathways hypothesized for distinct strategy uses (3a and 3b), including effortful counting strategies mediated by prefrontal working memory system (S1; dlPFC, dorsolateral prefrontal cortex), hippocampal-dependent episodic-like memory retrieval (S2; MTL), hippocampal-independent semantic-like memory retrieval (S3, PPC, posterior parietal cortex), we established feedforward NN models with back propagation. Twenty-two input nodes were used to provide distributed representations of two separate digits, and 3 hidden layers (20 units in each) were employed to simulate three pathways for multiple strategies used for arithmetic problem solving. Another 20 hidden units were also used to associate inputs and output patterns (18 units) coming from three pathways. Information from input to output is feedforward and each hidden layer is self-connected. The efficiency of each pathway is controlled by a set of input units to provide differential additional potentiation because units in each pathway were associated with large negative bias.



**Figure 3**

***Model training and testing***. Different models are trained to associate two number representations at input layer to one answer representation at output (44 problems in total). After training, differences across age groups (Child T1, Child T2, Adolescent and Adult) are manipulated by the proportion of pathways used to solve the 44 arithmetic problems. Based on data from previous study (Qin et al., 2014), from Child T1 to Adult, the proportion of using the most efficient S3 pathway gradually increases. The absolute output value on target unit is recorded to score the accuracy and latency of the model.

***Preliminary results and expected outcomes***. Only manipulating the pathway efficiency for different strategies, we are able to disentangle effects of using multiple strategies and other learning-related differences across age groups. We conducted some preliminary tests for the face validity of the models, and showed that with moderate amount of training, the efficiency of different pathways showed large effects on learning so different age groups show predicted behavioral discrepancies on accuracy (Figure 4). But with extensive training, the effect of efficiency diminishes yielding ceiling effects for all age groups on accuracy. We also predict age group difference on response latency that adolescent and adult groups should show faster responses than two child groups. Also, because adolescent and adult groups use the most efficient retrieval strategy dominantly, we predict that the trial-by-trial latency variance should be smaller than two child groups.



**Figure 4**

**Aim 2: Account for the neurobiological changes in neural activity patterns in different brain systems throughout development.**

***Neural network model rationale and specification***. The same NN models will be used.

***Neural activity analysis in the model and human subjects***. In the NN models, absolute activation values on hidden units will be recorded to be analogously compared to human fMRI data. We will extract the mean activation values in all three pathways for all four age groups to reveal the developmental changes. For the fMRI data on human subjects, data analysis was based on a total of 68 children, adolescents and adults participated in a published study in our lab; and 28 children with scanned twice as child T1 and T2 (Qin et al., 2014). Univariate general linear model analysis (GLM) will be used to identified group differences in activation, and multi-voxel pattern analysis (MVPA) will also conduced to search for brain regions showing high classifications of multi-group identities.

***Neural representational stability analysis***. In the NN models, Activation values of each testing problem of hidden units will be used to compute a problem-by-problem similarity matrix across all units within each pathway. Therefore, we will obtain three similarity matrices (S1-S3) for every age group as the interproblem representational stability measures. For the fMRI data on human subjects, a novel searchlight mapping method was used to obtain a whole-brain searchlight maps for a measure of interproblem representational stability in order to determine brain areas that exhibited similar developmental changes in interproblem representational stability as we observe in the models.

***Expected outcomes***. (1) Decrease in neural activity in S1 pathway of the model should be in line with decrease in neural activation in prefrontal-parietal working memory network across age groups, due to the less involvement of over counting strategies; (2) Increased neural activity in S2 pathway of the model from Child T1 to Child T2 should be observed as in human fMRI data (see Figure 2); and (3) Interproblem representational stability pattern across age groups in S2 pathway should be similar to MTL pattern in human subjects which show lower stability in children due to use a mix of counting and retrieval strategies.

**Aim 3: Explain the developmental transition of different strategy uses through learning. *Neural network model rationale***. In order to explain how children learn to use different strategies and what drives the shift of strategy uses, we need to extend the NN models to be a more ecological and neurally-faithful way by including other brain systems for perceptual, motor, and attentional processes as illustrated in Figure X. The NN models proposed in Aim 1 and 2 captures basic behavioral outcomes and neural activity patterns in visual ventral, MTL, prefrontal and parietal brain systems, but the interaction between different pathways and more subtle strategies are not accounted for. Furthermore, the learning of counting strategies also require inclusions of auditory inputs and outputs (before children learn to read) as well as motor systems such as for putting up the fingers. Therefore, we need to establish NN models with the complex interaction of these neurocognitive systems.

***Model establishment***. The initial step is to replace the S1 pathway in the previous NN models with a learning module based on addition strategy bank from Shrager and Siegler (1996). The NN model will then be trained to associate two numbers as addends on left and right with an answer while also be trained to adaptively choose multiple strategies from the addition strategy bank for matching an answer distribution output patterns based on empirical data (Siegler & Shrager, 1984). Thus, we can start asking how the learning of overt strategies relate to the shift to more efficient strategies. The next steps are adding in auditory and motor systems incrementally.

***Model manipulation and expected outcomes***. We aims to run the following simulations to reveal whether and how the NN models account for developmental transition of strategy uses.

For hypothesis 3A:

a) We will manipulate a confidence threshold for using retrieval strategy to simulate a U-shaped retrieval usage. We predict a u-shaped retrieval curve where retrieval is heavily used early in the simulation but rapidly extinguish, leading to an extended period of explicit strategy use, which will slowly tail off to retrieval again as an adult-like behavior.

b) With similar manipulation, we predict a tangential pattern for answer correctness: early usage of retrieval gives largely incorrect answers but as this early retrieval extinguishes, and overt strategies come to be employed, more correct answers will be seen gradually; and by the end, in more "adult" phase where retrieval is used, most answers are correct.

For hypothesis 3B:

c) We will manipulate the order of training strategies: the NN models are (1) trained to use both counting and retrieval strategies simultaneously; (2) trained to use just counting strategies first and then retrieval strategies; or (3) to use just retrieval strategies first and then counting strategies. We don’t have explicit predictions about the learning outcomes of the training order of strategies, but this will generate novel implications for learning and education for early math skills.

**Aim 4: Explore sources of individual differences in neural basis to explain typical and atypical development**.

***Model manipulation and expected outcomes***. Based on the NN models established in Aim 3, we will test the behavioral outcomes of different lesion procedures, such as random deleting units, adding random noises into connections between layers, and mean strength in connections (Dikina & McClelland, 2008; Plaut, 1997). We predict that early lesions in multiple processing systems in the network may lead to failure in establishing stable numerical representations, and thus affects the models’ capacity to solving arithmetic skills.