NeuroImage 216 (2020) 116464

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

The balance of rigor and reality in developmental neuroscience

Jessica F. Cantlon

Carnegie Mellon University, Pittsburgh, United States

ABSTRACT

The balance of realism and rigor in psychological research is essential to the development of rich and accurate theories about the developing brain. In the field of neuroimaging researchers have used predominantly controlled laboratory methods to decompose neural signals into meaningful functions but there is currently a push to integrate naturalistic conditions into neural measurement. Sometimes naturalistic methods are used to validate existing functional theories ecologically, and other times they are used in data-driven studies for exploration. This article assesses the value and risk of these approaches for understanding the developing brain.

During early development, children learn because they need to solve rich social and physical puzzles in their environments. The controlled study of cognitive development, however, often requires simplifying the environment. Traditional methods of psychology remove many elements of the natural environment in order to provide rigorous experimental control over a behavior of interest. Researchers choose the critical elements of behavior *a priori* and isolate them in a paradigm. This can create a problem wherein the motivating elements of the environment that trigger learning and behavior are missing, which can dampen the expression of children's behavior and cause theories of cognition to become narrowly staked around specific phenomena.

For example, in word learning paradigms, a child is commonly placed in a dark room with headphones and presented with a stream of isolated speech sounds that vary statistically in their sequential relations. The test of learning is whether children show contingent responses to the statistically varying patterns in the speech stream. But increasingly researchers have asked: What if the critical triggers for learning language are not contained within isolated words in an auditory stream? In fact, it has been known for some time that language learning in infants is a social transaction and depends on the goal of communication with the parent (Yurovsky, 2018; Tomasello and Carpenter, 2007) and that social, parent-related cues to parsing speech can even override infants' perception of statistics (Johnson and Jusczyk, 2001). For example, when motherese cues are available, statistical regularities in speech are not computed as readily by infants and instead they rely on motherese cues to parse words. Thus, although infants show some evidence of word learning during more stripped-down controlled laboratory tasks, the controlled paradigm may be missing key elements of the infants' computation, such as the parent's behavior. In this sense, relying exclusively on controlled paradigms could result in 'weaker than real life' learning observations in infants, or in theories that miss the primary inputs and mechanisms infants use to learn language in the real world (Smith et al., 2018).

Reliance on controlled tasks can sometimes lead to oversimplified theories. In numerical development, for example, significant focus has been placed on the task of two-alternative-forced-choice verbal and nonverbal number comparison. More and more those task representations seem to have a relatively limited causal role in learning mathematics for children. Instead, a heterogeneous constellation of domain-general and domain-specific representations can cause mathematics impairments (Bartelet et al., 2014). Field research shows that the to-be-explained phenomenon of mathematics learning is much broader 'in the wild' than what is represented in the cognitive literature and theory (Geary, 2011). One source of the problem here is that many controlled tasks of early numerical representation neutralize domain-general processes, making it impossible to observe dynamic interactions between domain-general and domain-specific representations over development.

The general point is that controlled laboratory design can make learning appear to be a neat process that draws on narrow subsets of representations and rules. The narrowness of the tasks employed can lead to theories that over-emphasize simplicity in mechanisms and representations at the expense of accuracy. To understand development, researchers must sometimes look broadly at the messy information space of the child and set aside *a priori* notions of what the eventual endpoint of development is. In developmental neuroscience, this would mean studying children's neural activity during thinking and reasoning about natural scenes, events, and problems.

Although most would agree that studying children under natural conditions is good, there is a tension between inferences made about cognition in the real world using unconstrained natural behavior (realism) versus those made by laboratory manipulation with

E-mail address: jcantlon@andrew.cmu.edu.

https://doi.org/10.1016/j.neuroimage.2019.116464
Received 8 July 2019; Received in revised form 14 November 2019; Accepted 11 December 2019
Available online 24 December 2019
1053-8119/© 2019 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

behavioristically controlled tests and pure contrasts of hypotheses (rigor). Developmental scientists have long been aware that flawed conclusions about children's cognition can result from too little or too much environmental control (Bronfrenbrenner, 1976; Greeno et al., 1984). Naturalistic studies have been criticized as "dustbowl empiricism" because they collect massive amounts of observations with minimal control, then analyze the data exhaustively, and report the outcomes with the 'theoretical clarity of a dust storm' (Bronfrenbrenner, 1976). On the other end, controlled laboratory-based methods have been criticized as "white room" research undertaking the "science of the strange behavior of children in strange situations with strange adults for the briefest possible periods of time." (Bronfrenbrenner, 1976). The criticism here is that laboratory methods exert strict control over the environment, at times yielding highly artificial paradigms and measurements without generalizability to the real world. These criticisms highlight the risk inherent in neglecting either rigor or realism.

Generalizability and ecological validity

Without studying mechanisms *in situ*, we do not know how they scale up to real world functions, or what percent of the variance in processing is explained by the isolated mechanisms we observe under controlled conditions – our models are incomplete and we lack the observations that are necessary to complete them.

Naturalistic methods play an essential role in completing theories. One way to employ naturalistic methods effectively is to develop a theory of neural processing using controlled functional decomposition in the laboratory and then test its ecological validity using naturalistic tasks. This approach is taken in cognitive development where researchers translate conclusions drawn in the lab to learning interventions in the field (eg., Dillon et al., 2017; McNeil and Uttal, 2009; Newcombe et al., 2009; Newcombe and Frick, 2010; Thomas et al., 2019).

In the neuroscience of numerical cognition, we and others have worked on the theory, originally derived from lesion studies of neuropsychological patients (Dehaene and Cohen, 1997) that a core neural substrate of numerical processing in the human brain is the intraparietal sulcus (IPS). We and others hypothesized that the role of the IPS in numerical processing is developmentally primitive because it has an evolutionary basis in visuospatial quantitative functions shared by primates (Ansari, 2008; Cantlon et al., 2006; Nieder and Miller, 2004). To test this, Alyssa Kersey and I conducted an fMRI adaptation study with 3- to 6-year-old children in which we presented children with a sequence of dot arrays that typically had the same number, color, and surface area; occasionally, deviant stimuli appeared in which these properties were altered (Kersey and Cantlon, 2017, Fig. 1A). Regions of the IPS adapted to the typical stimuli, and showed numerically-dependent release from adaptation during number deviants (Fig. 1A). Neural responses in the IPS to deviations in other dimensions (shape, color) were flat. The results implicate functional specialization for numerical representation, relative to other dimensions, in a region of the brain (the IPS) independently predicted to represent numerical quantity in young children.

Controlled designs such as those in Kersey and Cantlon (2017) maximize opportunities for contrast and afford mutually exclusive inferences about processes. However, the simple, ballistic, stripped-down nature of the stimulation and task could lead to an incorrect or oversimplified functional analysis of the brain. The controlled passive-viewing fMRI adaptation task neutralizes many aspects of active cognition as well as input from the environment. Despite showing that a region of cortex represents numerical value independently of other properties, it is unclear from the study whether that region is prominently engaged when children invoke mathematics concepts in the real world.

In an effort to validate our laboratory theory under more naturalistic conditions, we used fMRI to collect neural responses from young children as they watched 20 min of Sesame Street (Cantlon and Li, 2013, Fig. 1B). The episode contained segments about mathematics, reading, and other topics. A group of adults were imaged as they watched the same episode of Sesame Street to provide a measure of mature neural activity. We used inter-subject correlation to test the strength of the relation between children's neural timecourses and adults' timecourses, voxel-to-voxel

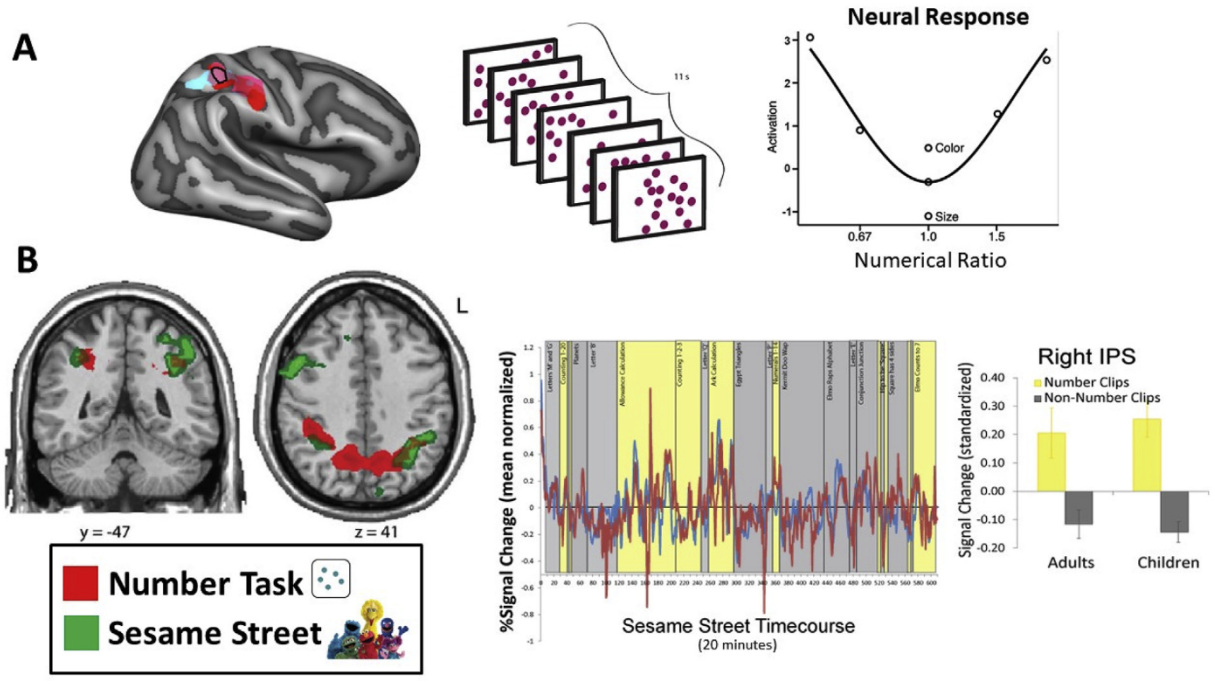

Fig. 1. A) In 3- to 6-year-old children regions of the IPS outlined in black show adaptation to a repeated numerical value and increased amplitude to deviant numerical values. The neural response to deviant values follows a ratio-based tuning curve (Kersey and Cantlon, 2017). B) In children ages 4 to 10 regions of the IPS respond to numerical stimuli in a controlled task (red) and to naturalistic number-related content in a Sesame Street video (green). During Sesame Street viewing, children's intersubject correlation to adults in the IPS is related to their mathematics ability (green). Percent signal change is higher in the IPS during 'number' clips (yellow) than other clips (gray; Cantlon and Li, 2013).

throughout the brain, yielding a metric of neural maturity for each child, at each voxel in the brain. A key outcome was that the maturity of children's neural timecourses in the IPS predicted their mathematics ability (independently of verbal ability), measured by their scores on a standardized mathematics test (Fig. 1B). Moreover, children's neural responses in the IPS were higher during mathematics content than during the non-numerical content of Sesame Street (Fig. 1B). IPS regions that showed sensitivity to mathematical processing during Sesame Street also showed sensitivity to numerosity in controlled tasks tested on the same children. Thus we were able to validate our functional theory about the developing brain with an ecologically relevant task – watching Sesame Street.

Richardson et al. (2018) used a similar approach with socially complex Pixar movies to validate and extend their functional theory about the role of the temporo-parietal junction network in the development of theory of mind (TOM). The results showed functionally specific neural responses in the predicted TOM network when 3- to 12-year-old children viewed movie content invoking theory of mind, and that these neural responses were dissociable from those related to content about bodies. The relevance of their functional theory, derived from many controlled laboratory studies of child development, was thus extended to the real world that children experience.

Naturalistic studies such as these are essential to the science of neural functions because they take the additional step of proving the validity and generalizability of a theory in the real world. Under this approach, the "white room" experiment is the first step in a scientific enquiry to develop hypotheses, and the naturalistic study represents the generalization phase to show that a mechanism functions in 'real life' as surmised. Real world generalization is critical because erroneous or incomplete conclusions can result if neural theories are not tested in natural conditions – even for simple processes and predictions (eg., Snow et al., 2011; Marini et al., 2019). For example, Snow and colleagues have shown significant differences in human visual perception and patterns of neural activation during object processing if those objects are presented as real-world tangible objects instead of the more commonly used 2D images of those objects.

Going further, Cantlon and Li (2013) and Richardson et al. (2018) showed that children's neural activity during naturalistic tasks is not fully predicted by activation from laboratory tasks. Cantlon and Li (2013) showed that children's neural responses to real-world mathematics stimuli were better predictors of their mathematics achievement scores than neural responses from a traditional laboratory task of number comparison. In the naturalistic task of Richardson et al. (2018), the theory of mind network extended beyond the temporoparietal junction and medial prefrontal cortex to the left angular gyrus and middle frontal gyrus, and showed less right temporal lobe activation than predicted by traditional tasks (Jacoby et al., 2016). The functional consequences of these differences are unclear and, moreover, it is unknown what percent of the variance in processing is accounted for by the predicted regions versus those revealed unexpectedly. These findings illustrate how the to-be-explained phenomena of real world neural processing go beyond what the 'white room' approach can explain. A risk of using naturalistic studies too narrowly, only to validate functional theories from the laboratory, is that those theories might contain oversimplifications of neural functions. Naturalistic studies that are too bound to particular functional theories could overlook meaningful signals of brain function that were not predicted by the theory and emerge unexpectedly under naturalistic conditions.

A further advantage of naturalistic stimulation is that it can yield distinct neural signals compared to controlled tasks even within a single network (Redcay et al., 2010). Neural variance, power spectral density, reliability, and multi-scale entropy differ between naturalistic and controlled tasks (Vanderwal et al., 2019). Distinctions between naturalistic and laboratory neural patterns likely have important functional implications. An untested assumption in controlled designs is that the real world differs from the controlled task only quantitatively or in "complexity" but conclusions about neural phenomena in controlled tasks might not scale up so simply. As described earlier, overly simplistic paradigms sometimes neutralize and remove neural processes that are critical to human cognition and development. In order to observe the real world neural phenomena we are trying to explain, it is important to examine the developing brain under conditions of functional complexity.

Exploration & data-driven discovery

Naturalistic methods generate discovery through exploration, and are critical for advancing methods and theory. Field studies using fNIRS and EEG in the classroom, and fMRI studies have produced in-principle demonstrations of feasibility for collecting neural measures in natural scenarios, and have generated new approaches to signal analysis (eg., Cohen et al., 2017; Dikker et al., 2017; Mason and Just, 2016; Yücel et al., 2017; Zaki and Ochsner, 2009). Already, discoveries of new types of neural patterns have resulted from data-driven explorations of higher-order neural activity from naturalistic paradigms.

Emergent properties such as neural synchrony or coherence, frequency oscillations, graph networks, intersubject correlation, and hierarchical timescales were initially observed in global analyses of neural activity during naturalistic thought (Berger, 1929; Bassett and Bullmore, 2006; Sporns et al., 2004). These higher-order neural properties are sensitive to semantics, episodic memory, and the unfolding of meaning over time. For example, Hasson et al. (2015) showed that higher-order temporal patterns of neural processing during naturalistic thought are hierarchical and functionally related to human memory. These observations occurred because researchers explored global patterns in neural data from naturalistic scenarios.

In developmental neuroimaging, data-driven analyses of child-specific neural patterns could yield new insights into the unique features of the child's brain. Natural viewing and listening studies have revealed that children and adults exhibit systematic temporal patterns of activation that are correlated from subject to subject, across large swaths of cortex (Cantlon and Li, 2013; Lerner, Scherf, et al., in press; Richardson et al., 2018). In children's and adults' neural responses to educational videos about reading and mathematics, we found that some temporal patterns are more similar among adults than between children and adults, and other patterns are more similar among children than among adults or between children and adults (Fig. 2A; Kersey, Wakim, Li and Cantlon, in press). That is, there are child-specific patterns of neural activation that are systematic among children, and distinct from adult patterns. These child-unique activations sometimes fall outside regions predicted for reading or mathematics by controlled tasks and contrasts (Fig. 2B). This finding highlights an advantage of studying complex functions like early reading and mathematics naturalistically and using data-driven methods: the potential to observe a broader range of neural patterns and functions, some of which may be unique to children's brains and difficult to elicit with *a priori* stimulus choices in controlled task designs.

The risk of data-driven exploratory approaches, however, is that they could leave us without coherent theories of brain function. Exploratory methods can yield interpretation without understanding. In this sense, exploratory methods in neuroscience are statistical but not scientific – they are missing an explanatory theory. For example, an exciting recent study observed correlated EEG signals among school children during simultaneous recordings in a natural classroom and showed that the degree of student-to-student correlation predicted class engagement (Dikker et al., 2017). The mechanisms mediating EEG synchrony and classroom engagement in these students are unclear, but the authors hypothesize a mediating role for joint attention. This is an example of the tradeoff between rigor and discovery wherein important new observations of neural patterns are derived from exploratory measures with limited ability to determine the underlying or causal mechanisms on the one hand, but with a strong ability to generate new testable functional theories on the other hand.

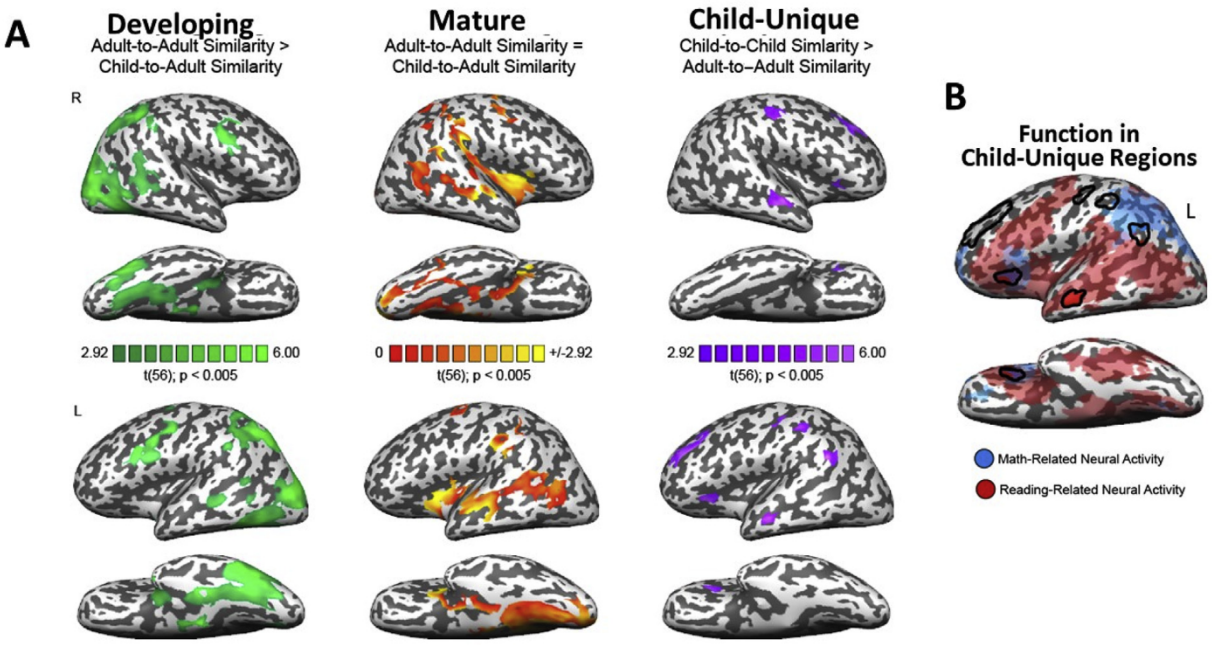

Fig. 2. A) Intersubject correlation measures provide metrics of developing (adult-adult > child-adult), mature (adult-adult = child-adult), and child-unique (child-child > adult-adult) patterns of neural activation from natural viewing of reading and mathematics videos in 4- to 8-year-olds (Kersey, Wakim, Li and Cantlon, in press). B) Child-unique regions (black outline) during naturalistic educational videos are not fully predicted by activation during simple tasks or contrasts with reading (red) and mathematics (blue).

Conclusion

Exploration and ecological validity are critical components of good science, and key advantages of naturalistic methods. In order to ensure rigor and realism in naturalistic studies, and to avoid the pitfalls of neglecting either of them, an ideal entry point is a comparison approach in which lab-controlled and naturalistic analyses are compared directly, in a single sample to observe the divergence between laboratory and naturalistic effects. Such a comparison would reveal the degree to which naturalistic neural patterns are explained by existing functional theories, and will expose those patterns that remain to be explained. Divergence between laboratory and naturalistic effects could point to functions typically subtracted out or neutralized in controlled tasks but not in naturalistic tasks, or they could point to regions that serve unpredicted functions in naturalistic tasks. Hypotheses and theories that emerge from those divergent neural patterns will inform the design of subsequent experiments and analyses. The combination of rigor and realism could create a more complete understanding of the human brain and its development than we could achieve with controlled tasks alone.

Author contributions

JC wrote the manuscript.

Acknowledgements

This work was funded by the National Institute of Health (R01 HD064636) and the National Science Foundation (DRL 1459625) in the United States.

References

Ansari, D., 2008. Effects of development and enculturation on number representation in the brain. Nat. Rev. Neurosci. 9 (4), 278.
Bartelet, D., Ansari, D., Vaessen, A., Blomert, L., 2014. Cognitive subtypes of mathematics learning difficulties in primary education. Res. Dev. Disabil. 35 (3), 657–670.
Bassett, D.S., Bullmore, E.D., 2006. Small-world brain networks. The Neuroscientist 12 (6), 512–523.
Berger, H., 1929. Über das Elektrenkephalogramm des Menschen. Archiv für Psychiatrie und Nervenkrankheiten 87 (1), 527–570.
Bronfrenbrenner, U., 1976. The experimental ecology of education. In: Paper Presented at the American Educational Research Association, April 19-23, 1976. San Francisco, CA.
Cantlon, J.F., Brannon, E.M., Carter, E.J., Pelphrey, K.A., 2006. Functional imaging of numerical processing in adults and 4-y-old children. PLoS Biol. 4 (5), e125.
Cantlon, J.F., Li, R., 2013. Neural activity during natural viewing of Sesame Street statistically predicts test scores in early childhood. PLoS Biol. 11 (1), e1001462.
Cohen, J.D., Daw, N., Engelhardt, B., Hasson, U., Li, K., Niv, Y., et al., 2017. Computational approaches to fMRI analysis. Nat. Neurosci. 20 (3), 304.
Dehaene, S., Cohen, L., 1997. Cerebral pathways for calculation: double dissociation between rote verbal and quantitative knowledge of arithmetic. Cortex 33 (2), 219–250.
Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., et al., 2017. Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. Curr. Biol. 27 (9), 1375–1380.
Dillon, M.R., Kannan, H., Dean, J.T., Spelke, E.S., Duflo, E., 2017. Cognitive science in the field: a preschool intervention durably enhances intuitive but not formal mathematics. Science 357 (6346), 47–55.
Geary, D.C., 2011. Cognitive predictors of achievement growth in mathematics: a 5-year longitudinal study. Dev. Psychol. 47 (6), 1539.
Greeno, J.G., Riley, M.S., Gelman, R., 1984. Conceptual competence and children's counting. Cogn. Psychol. 16 (1), 94–143.
Hasson, U., Chen, J., Honey, C.J., 2015. Hierarchical process memory: memory as an integral component of information processing. Trends Cogn. Sci. 19 (6), 304–313.
Jacoby, N., Bruneau, E., Koster-Hale, J., Saxe, R., 2016. Localizing Pain Matrix and Theory of Mind networks with both verbal and non-verbal stimuli. Neuroimage 126, 39–48.
Johnson, E.K., Jusczyk, P.W., 2001. Word segmentation by 8-month-olds: when speech cues count more than statistics. J. Mem. Lang. 44 (4), 548–567.
Kersey, A.J., Cantlon, J.F., 2017. Neural tuning to numerosity relates to perceptual tuning in 3–6-year-old children. J. Neurosci. 37 (3), 512–522.
Kersey, A. J., Wakim, K., Li, R., & Cantlon, J. F. (in press). Developing, mature, and unique functions of the child's brain in reading and mathematics. Developmental Cognitive Neuroscience.
Lerner, Y., Scherf, K. S., Katkov, M., Hasson, U., and Behrmann, M. (in press). Age-related Changes in Neural Networks Supporting Complex Visual and Social Processing in Adolescence.
Marini, F., Breeding, K.A., Snow, J.C., 2019. Distinct visuo-motor brain dynamics for real-world objects versus planar images. Neuroimage 195, 232–242.
Mason, R.A., Just, M.A., 2016. Neural representations of physics concepts. Psychol. Sci. 27 (6), 904–913.
McNeil, N.M., Uttal, D.H., 2009. Rethinking the use of concrete materials in learning: perspectives from development and education. Child development perspectives 3 (3), 137–139.
Newcombe, N.S., Ambady, N., Eccles, J., Gomez, L., Klahr, D., Linn, M., et al., 2009. Psychology's role in mathematics and science education. Am. Psychol. 64 (6), 538.

Newcombe, N.S., Frick, A., 2010. Early education for spatial intelligence: why, what, and how. Mind, Brain, and Education 4 (3), 102–111.
Nieder, A., Miller, E.K., 2004. A parieto-frontal network for visual numerical information in the monkey. Proc. Natl. Acad. Sci. 101 (19), 7457–7462.
Redcay, E., Dodell-Feder, D., Pearrow, M.J., Mavros, P.L., Kleiner, M., Gabrieli, J.D., Saxe, R., 2010. Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. Neuroimage 50 (4), 1639–1647.
Richardson, H., Lisandrelli, G., Riobueno-Naylor, A., Saxe, R., 2018. Development of the social brain from age three to twelve years. Nat. Commun. 9 (1), 1027.
Smith, L.B., Jayaraman, S., Clerkin, E., Yu, C., 2018. The developing infant creates a curriculum for statistical learning. Trends Cogn. Sci. 22 (4), 325–336.
Snow, J.C., Pettypiece, C.E., McAdam, T.D., McLean, A.D., Stroman, P.W., Goodale, M.A., Culham, J.C., 2011. Bringing the real world into the fMRI scanner: repetition effects for pictures versus real objects. Sci. Rep. 1, 130.
Sporns, O., Chialvo, D.R., Kaiser, M., Hilgetag, C.C., 2004. Organization, development and function of complex brain networks. Trends Cogn. Sci. 8 (9), 418–425.
Thomas, M.S., Ansari, D., Knowland, V.C., 2019. Annual research review: educational neuroscience: progress and prospects. JCPP (J. Child Psychol. Psychiatry) 60 (4), 477–492.
Tomasello, M., Carpenter, M., 2007. Shared intentionality. Dev. Sci. 10 (1), 121–125.
Vanderwal, T., Eilbott, J., Castellanos, F.X., 2019. Movies in the magnet: Naturalistic paradigms in developmental functional neuroimaging. Developmental Cognitive Neuroscience 35 (100600).
Yücel, M.A., Selb, J.J., Huppert, T.J., Franceschini, M.A., Boas, D.A., 2017. Functional near infrared spectroscopy: enabling routine functional brain imaging. Current opinion in biomedical engineering 4, 78–86.
Yurovsky, D., 2018. A communicative approach to early word learning. New Ideas in Psychology 50, 73–79.
Zaki, J., Ochsner, K., 2009. The need for a cognitive neuroscience of naturalistic social cognition. Ann. N. Y. Acad. Sci. 1167 (1), 16–30.