White- and Grey-Matter Damage Differentially Impair Learning and Generalization in a Computational Model of the Raven Matrices Task

Vincent G. Berthiaume (Vincent.Berthiaume@McGill.ca)
Department of Psychology, McGill University, 1205 Dr. Penfield Avenue
Montréal, QC H3A 1B1 Canada

Thomas R. Shultz (Thomas.Shultz@McGill.ca)
Department of Psychology and School of Computer Science, McGill University, 1205 Dr. Penfield Avenue
Montréal, QC H3A 1B1 Canada

Olaf Dammann (ODammann@TuftsMedicalCenter.org)
Division of Newborn Medicine, Floating Hospital for Children at Tufts Medical Center, 800 Washington Street, Box 854
Boston, MA 02111 USA

Abstract

Many preterm neonates have white-matter damage (WMD, damaged connections between neurons) and grey matter-damage (GMD, dead neurons). These children are known to have lower IQs than their full-term peers, yet the mechanisms underlying this association are poorly understood. We designed a developmental connectionist model of the Raven Matrices IQ task in which (1) all neurons had intact output, simulating normal development, or (2) half the neurons had noisy output, simulating noisy transmission or WMD, or (3) half the neurons had no output, simulating cell death or GMD. We found that damage increased task error. Further, WMD was worse than GMD overall, yet GMD was at once worse for generalization problems not given in training and better for training problems. Our model is the first to simulate an effect of perinatal brain damage on a cognitive task, and predicts that different types of brain damage may lead to different cognitive impairments.

Keywords: White-matter damage; cortical damage; preterm birth; Raven Matrices; IQ; connectionism; learning.

Background

In 2007, 12.7% of all births in the United States were preterm, an increase of over 2% since 1990 (Heron et al., 2009). This increase inevitably exacerbates family distress and healthcare costs, as children born preterm present many cognitive and developmental impairments compared to their full-term peers, including lower IQ scores (Bhutta, Cleves, Casey, Cradock, & Anand, 2002). The severity of preterm children's cognitive deficits appears to be correlated with brain abnormalities, e.g., reduced volume in specific brain regions (Peterson et al., 2000), which may result from abnormal development following perinatal brain damage (Robinson, 2005). Indeed, preterm neonates have immature brains that are likely to suffer damage from prematurity-associated adverse exposures before and after birth.

Perinatal brain damage can occur in either of the two major macroscopically distinct areas of the brain, the white (Dyet et al., 2006) and grey matter (Burd et al., 2009). White matter is made up of myelinated axons connecting neuronal regions and is the matter principally damaged in preterm brains (Leviton & Paneth, 1990). By contrast, grey matter consists of neuronal cell bodies and its damage is usually more constrained in the preterm brain (Billiards, Pierson, Haynes, Folkerth, & Kinney, 2006). Although the association between cognitive impairments and brain damage is well known in the pediatric community, not much is known about either the general mechanisms underlying the association (Counsell et al., 2008), or more specifically, about how damage to white or grey matter may potentially affect cognitive function differentially. Although a previous computational model indicated that white-matter damage may be worse than grey-matter damage for synaptic recovery (Follett, Roth, Follett, & Dammann, 2009), that model did not implement any cognitive task and thus did not inform us about the effect of damage on cognition.

In order to explore how white- and grey-matter damage may affect cognitive ability, we designed a computational developmental model of a popular IQ task, the Raven Matrices, and incorporated white- and grey-matter damage in the model to assess their effects on task performance.

Computational Developmental Algorithm

Sibling-Descendent Cascade-Correlation (SDCC, Baluja & Fahlman, 1994) is a supervised-learning, artificial-neural-network algorithm which benefits from fast and powerful learning and implements some psychologically- and neurologically-plausible mechanisms (Shultz, 2006; Shultz, Mysore, & Quartz, 2007). Its developmental or constructive aspect comes from the fact that networks initially have only input and output units (fully interconnected with random weights), but develop by recruiting hidden units, as required to reduce error in training.

Training includes output and input phases. Networks are first given training patterns (input and target patterns), and training enters the output phase, in which the algorithm reduces output error, the discrepancy between output activation (initially random) and the target patterns. If the algorithm cannot bring error lower than the Score Threshold (ST) parameter, left at its default value of .4 for all training patterns, training switches to the input phase. In the input phase, the network selects the one hidden unit, out of a pool of 8 randomly-initialized candidate recruits, that correlates most with output error. This selected unit is integrated into the network and training switches back to output phase. Training usually stops as soon as error for each training pattern drops below the ST. However, in order to have consistent amount of training across all types of networks, we imposed here a training limit of 14 hidden units and 2500 epochs, based on the average training cost of an independent, undamaged sample of 100 networks.

At the end of training, networks are tested by freezing connection weights (so that networks do not learn during testing), and measuring output error on testing patterns.

Raven Matrices task

The Raven Matrices task consists of a series of problems, in which subjects have to study a 3-by-3 matrix, and chose amongst 8 alternatives the figure that best fits the empty spot in the matrix (Figure 1).

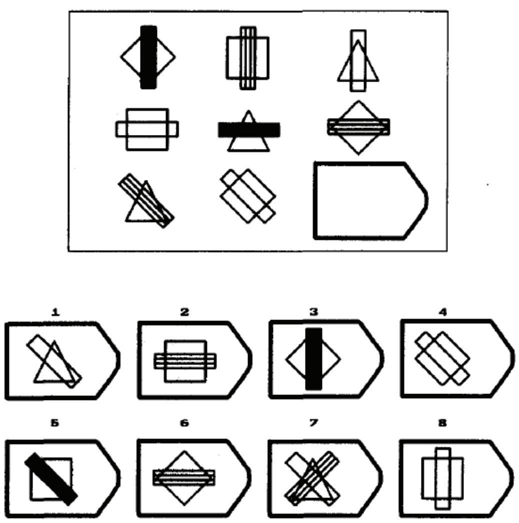

Figure 1. An example Raven problem. Copyright © 1990 by the American Psychological Association. Reproduced with permission from Carpenter, Just, and Shell (1990). The use of APA information does not imply endorsement by APA.

There are four rules (Carpenter et al., 1990) for predicting the missing figure. In the constant-in-a row rule, a figure feature is constant across rows. For example, the narrow rectangle in Figure 1 is always vertical in the first row, horizontal in the second, and diagonal in the third. In the distribution-of-three rule, a feature is distributed amongst the figures in a row, e.g., the narrow rectangle is either black, striped, or transparent in each column in Figure 1. If one of the three features is absent, the distribution-of-three rule can also cover a distribution-of-two-values rule, sometimes considered as a separate rule. In the quantitative-pairwise-progression rule, figure attributes (such as small squares in a grid) increment or decrement between adjacent columns. In the addition and subtraction rules, a figure feature from column 1 is added to or subtracted from a figure in column 2 to produce a third figure in column 3.

Methods

We used SDCC to train and test undamaged networks on the Raven Matrices task. We next incorporated damage in two different groups of networks by either randomizing (white-matter damage) or blocking (grey-matter damage) the output activation of approximately half the networks' neurons.

Undamaged Training and Testing

A first group of 100 undamaged networks were trained and tested on Raven task problems that each implemented one of the four rules identified by Carpenter and colleagues (1990). Performance was evaluated on problems that networks knew about, and on novel problems, a technique somewhat similar to some psychological studies using the Raven task (e.g., Skuy et al., 2002).

Networks had eight inputs corresponding to the eight figures constituting a Raven problem, and one output corresponding to the missing ninth figure. Inputs and outputs used linear activation functions to cover the range of possible input and output values (see below). In order to compare network performance on known and novel data, two datasets of equal size were constructed: the training and generalization sets. Figure 2 illustrates an example Raven problem coded for training and generalization patterns.

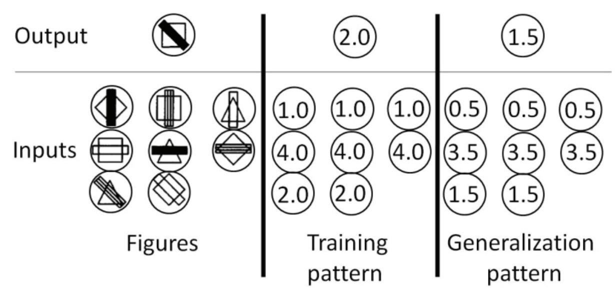

Figure 2. A Raven problem represented in figures and as a training pattern, and its derived generalization pattern.

The left-most panel of Figure 2 shows the example figures, and the middle panel shows how the figures may be coded as a training pattern. For each training pattern, selected features were coded by integers (chosen at random between 1 and 4 from a uniform distribution) that represented the figure feature relevant to the problem rule. Each training pattern implemented one of the 4 rules identified by Carpenter and colleagues, (1990). For instance, in this constant-in-a-row example problem, 1.0 represents a vertical bar, 4.0 an horizontal one, and 2.0 a diagonal one.

The right panel shows a generalization pattern, obtained by subtracting .5 from every value of the example training pattern. Following previous practice (Dandurand, Berthiaume, & Shultz, 2007), generalization patterns were all obtained using this calculation (although in feature-addition and -subtraction problems, .5 was only subtracted from numbers in the first two columns, because the third value depended on the first two). Other types of problems were coded similarly. Distribution-of-three problems had one of three numbers appear in each column. Quantitative-pairwise-progression problems were represented by an increment or decrement of numbers across adjacent columns. Addition and subtraction problems had a number from the second column added to or subtracted from the number the first column, to produce the third column number (in subtraction problems, the first column value was always bigger than in the second column, to ensure positive values in the third column). The range of input and output values was [.5, 8.0], where [5.0, 8.0] were only present when due to the addition of other features, i.e., [1.0, 4.0] for the training set and [.5, 3.5] for the generalization set.

Training and generalization sets each included 20 examples of each of the 5 types of Raven problems (feature-addition and-subtraction were considered 2 different types), for a total of 100 problems. Each dataset was created by sampling randomly, with possible repetitions of rows and problems, through the possible permutations of the 4 feature values, so that no network had identical training or testing. In test, after training, we calculated mean squared output error for both training and generalization datasets.

Damaged Training and Testing on the Raven task

Two other groups of 100 networks were trained and tested as described above, except that they were damaged by either randomly reducing (white-matter damage) or blocking (grey-matter damage) the output activation of some of their neurons. Damaged neurons were selected randomly for each network, and half of the input neurons and half of the candidate hidden neurons were damaged. There is nothing special about impairing half the neurons, we selected that proportion as a starting point for our experiments. Networks were free to recruit or not recruit impaired hidden neurons, so as to simulate more naturally perinatal brain damage, i.e., prior to learning and performing on tasks. The output neuron was not damaged, in order to insure a fairer comparison of white- and grey-matter damage (a grey-matter-damaged output would prevent any network output).

White-matter damage. White-matter damage is often observed as abnormal white-matter signal and abnormal axonal myelination (Counsell et al., 2006). A reduction in white-matter signal may be due to noisy or leaky axonal transmissions in which abnormal axonal myelination causes action potentials to be lost. To model this leaky transmission we subtracted a different random value from the activation value of impaired neurons each time an activation value was calculated, as in:

$$A_r = Activation - [Activation \times RandomValue(0,1)]$$

where A_r is the reduced random activation, *Activation* is the undamaged activation and *RandomValue(0,1)* is a value chosen randomly from a [0, 1] uniform distribution.

Grey-matter damage. Grey-matter damage can be considered as cell death, leading to a complete loss of signal (e.g., Follett et al., 2009). It was therefore modeled by reducing the activation values of each impaired neuron to 0.

Results

After training, we performed a two-way between networks analysis of variance (ANOVA) in order to compare the effects of dataset (training, generalization) and damage type (undamaged, grey-matter, white-matter) on mean output error. The main effects of dataset and damage type, as well as the dataset by damage type interaction, were all significant. Figure 3 shows mean output error for the different datasets and damage types.

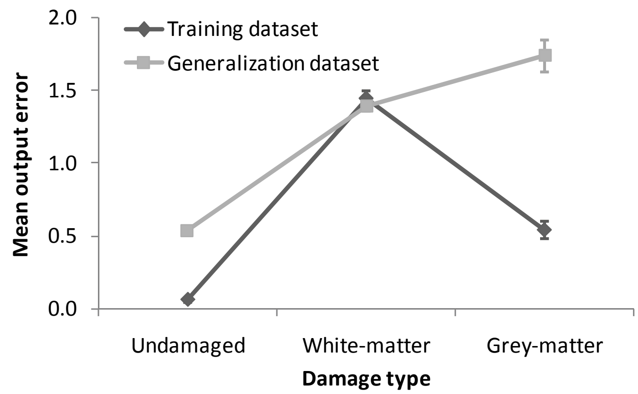

Figure 3. Mean output error and SE bars for the different datasets and damage types. Due to low variation, error bars in the undamaged condition are not clearly visible.

First, error was higher for the generalization, M = 1.22, SD = .83, than for the training set, M = .68, SD = .73, $F(1, 594) = 139$, $p < .001$. It is common for networks to perform better on problems on which they have been trained.

Second, the significant effect of damage type, $F(2,594) = 213$, $p < .001$, was explored using Bonferroni post-hoc tests. Error was significantly lower for the undamaged condition, M = .30, SD = .31, than for either the white-matter, M = 1.42, SD = .42, or grey-matter damage condition, M = 1.14, SD = 1.04, ps < .001. Further, error was significantly lower for grey- than for white-matter damage networks, p = .001.

Third, to explore the significant dataset by damage type interaction, $F(2,594) = 62$, $p < .001$, we analyzed mean network error for each level of the factor dataset (training, generalization), using one-way ANOVAs with damage type (undamaged, grey-matter, white-matter). For the training set, the effect of damage type was significant, $F(2, 297) = 250$, $p < .001$, and Bonferroni post-hoc tests revealed that error was significantly lower for the undamaged condition, M = .06, SD = .12, than for either grey-, M = .54, SD =.58, or white-matter damage, M = 1.44, SD = .49, with error being significantly lower error for the grey- than the white-matter damage, ps < .001. For the generalization set, the effect of damage type was also significant, $F(2, 297) = 87$, p

< .001, and error was still significantly lower for undamaged, M = .54, SD = .24 than for either grey, M = 1.74, SD = 1.06, or white-matter damage, M = 1.39, SD = .34, ps < .001. However, this time error was significantly lower for *white*- than for grey-matter damage, p = .001.

Discussion

We modeled undamaged, white-matter-damage and grey-matter-damage performance on the Raven Matrices task. Of the three conditions, white-matter damage produced highest error. However, the damage type by dataset interaction revealed that compared to white-matter damage, grey-matter damage produced at once higher error for generalization problems not seen in training, and lower error for problems seen in training. To our knowledge, our computational model is the first to demonstrate an association between white- and grey-matter damage and cognitive impairment.

White- worse than grey-matter damage overall

Why was white-matter damage, i.e., noisy reduced axonal signal, overall worse than grey-matter damage, i.e., no axonal signal at all? This perhaps unexpected result may be due to white-matter damage varying in time. That is, white-matter damaged neurons had different noise values every time activation values were calculated, whereas grey-matter damaged activation values were constantly null. White-matter damage networks thus had to deal with changing information, whereas grey-matter damage networks—although missing considerable information—could adapt better to their damage because at least it was constant.

In their computational model of synaptic recovery, Follett and colleagues (2009) also reported a worse effect of white- compared to grey matter-damage, but their model did not test cognitive impairment. Our model adds to their findings by indicating that white- may be worse than grey-matter damage for learning and performing on cognitive tasks. Our results may thus provide insights into the mechanisms underlying the association between damaged and/or reduced white-matter structure and reduced cognitive abilities in preterm children (Skranes et al., 2007), full-term children (Schmithorst, Wilke, Dardzinski, & Holland, 2005) and normal, age-related cognitive decline (Charlton et al., 2006).

Damage type and dataset interaction

Even though error was overall larger for white- than grey-matter damage, grey-matter damage produced larger error on generalization problems, i.e., problems not used in training. Our model thus predicts that different types of perinatal brain damage may be associated with different types of cognitive impairment. It is however difficult to compare our predictions with findings from the preterm literature as not much is currently known about white- versus grey-matter damage in cognitive development (Dammann, Kuban, & Leviton, 2002), and because preterm children with grey-matter damage generally also have white-matter damage, (Pierson et al., 2007). Further, the association between preterm perinatal grey-matter damage and cognitive impairments has not yet been studied directly.

Why different effects?

Interestingly, our further simulations (not reported here) indicate that the differential effects of white and grey-matter damage still hold when the imposed training limit is either doubled or cut in half, when using generalization patterns drawn from the same distribution as training patterns, as well as on the continuous XOR benchmark problem. In continuous XOR there are 2 inputs, each varying between [-.5, .5] and the output is 1 when inputs indicate a point in either the first or third quadrant, and zero in the other two quadrants. The interaction thus seems to be robust to changing the training length and the task.

Insight into our findings may be achieved by analyzing other computational studies. We implemented white-matter damage by randomly reducing the output activations of damaged neurons. Such manipulations resemble injection of noise in neural-network simulations, which was previously found to improve generalization. For instance, Jim, Giles, and Horne (1996) found improved generalization on a dual-parity problem and a randomly generated six-state problem by adding noise to the connection weights of their networks. Unsworth and Coghill (2006) also found improved generalization in their multilayer perception networks, designed to recognize partially obscured human movement, but this time by injecting noise in the training data.

Adding noise can thus improve generalization, perhaps explaining better generalization for white than grey-matter damage. Generalization was however worse for white-matter damage than for *undamaged* networks. This may be due to very high training error in white-matter damage (more than four times higher than for undamaged networks). Indeed, networks' generalization is limited by the quality of their learning. Because white-matter damaged networks had high training error, their overall generalization error was also high. Further, Figure 3 reveals white-matter damage to be the only condition in which error is *not* higher for generalization than training problems (in fact it appears to be slightly *lower* for generalization), which suggests some improved generalization in white-matter damaged networks.

Our implementation of white matter damage differed from the previous noisy simulations. Compared to others who injected noise in either connection weights (e.g., Jim et al., 1996) or in the training data (e.g., Unsworth & Coghill, 2006), we injected noise at the level of neurons' output activations, to simulate impaired axonal transmission. Further, whereas others have used absolute, small noise values, e.g., between [0, 2] (Jim et al., 1996), we used proportional, large noise values that varied between 0% and 100% of neurons' output activations. Thus our noise values varied between [.5, 8.0] due to the range of possible values in the input patterns. Therefore, white-matter damage may have produced large error due to the large noise values.

We implemented grey-matter damage by blocking the output of damaged neurons, simulating cell death and no axonal transmission. This manipulation resembles neuronal pruning, usually used to increase generalization in neural networks (Reed, 1993). However, pruning algorithms usually select smaller, less important connection weights to be deleted (LeCun, Denker, & Solla, 1990). The idea is that large networks may use their extra connections to encode some of the specifics of the training data. Pruning algorithms thus usually remove smaller weights, in the hope that the remaining, larger connection weights better encode the pattern underlying the data. By contrast to these connection pruning techniques, our networks had whole neurons damaged and these neurons were chosen at random, without regards to whether they were important or not for task performance. Removing potentially critical neurons and connections, as opposed to non-important ones, may explain why grey-matter damage worsened generalization rather than improve it like pruning algorithms.

It is still unclear why training error was lower for grey- than for white-matter damage. This result may reflect the intuition that learning may be easier when missing some information compared to when having wrong information. For instance, Eggert, Ladda, and Straube (2009) found that subjects were better at predicting the trajectory of dots on a screen if *no* aiding cues were provided compared to when both correct and misleading cues were provided. In the case of grey-matter damage, networks apparently learned training problems without the missing input neurons. By contrast, networks with white-matter damage received information from all their input neurons, including some misleading, noisy information which may have made it difficult to learn.

Future directions

We simulated the Raven task by assigning random values to the main features of the matrix figures, and arranging these values in problems following any of the four Raven rules (Carpenter et al., 1990). By contrast, real Raven matrix figures often contain several features which vary along several rules, and thus human subjects have to find which of the features are relevant to which rules. Future simulations may more closely match the task, e.g., by using vectors or sub-matrices to encode all the figures' features. However, because networks still had to figure out the four rules only from the pattern of inputs, we consider our task to still be quite challenging. An indication of this difficulty may lie in the fact that many hidden neurons, i.e., 14 on average, were required by undamaged networks to learn the task. Further analyses may also use the number of problems solved correctly rather than using the usual output error measure. We could thus study whether white- and grey-matter damage also have differential effects on the number of problems solved, and assess the order in which networks succeed at different types of problems as they develop.

We implemented white- and grey-matter damage by impairing half of the neurons in damaged networks (excluding the single output neuron), and damage was static, i.e., a given damaged neuron stayed damaged for the whole simulation. However, because the infant brain is very plastic, perinatal brain damage may interact in a complex way with the child's later development. Future work may consider developmental damage, e.g., punctual damage only at the beginning rather than throughout the simulation, or that is more closely related to the networks' hidden neuron recruitment. For instance, an area often damaged in the preterm brain is the germinal matrix, which is responsible for generating cortical neurons. Because white-matter damage is associated with damage to neurons migrating from the germinal matrix (Leviton & Gressens, 2007), future simulations may more closely simulate perinatal brain damage by directly impairing the hidden neuron recruitment process in SDCC, rather than letting networks decide whether to recruit damaged or undamaged neurons. We may also compare networks with different proportions of both white- and grey-matter damage.

Summary

Our computational model explored the potential link between brain damage and cognitive impairments in preterm children. White-matter damage produced overall higher task error, but grey-matter damage produced higher error on generalization problems, not seen in training. Our results thus predict that different types of brain damage may lead to different types of cognitive impairments. Future psychological work may test this prediction, e.g., by having white- and grey-matter damage populations trained on Raven problems and tested on novel problems (perhaps using a procedure similar to Skuy et al., 2002). Insights gained into the mechanisms underlying the association between perinatal brain damage and cognitive impairment may lead to more effective treatment for survivors of prematurity and help alleviate this aggravating problem.

Acknowledgments

We thank Kristine H. Onishi for insightful discussion and helpful comments. This research was supported by a scholarship to V.G.B. and a grant to T.R.S. from the Natural Sciences and Engineering Research Council of Canada, and by grants to O.D. by the Richard Saltonstall Charitable Foundation, the National Institutes of Health (R21Y019253), and the European Commission (LSHM-CT-2006-036534, HEALTH-F2-2009-241778.

References

Baluja, S., & Fahlman, S. E. (1994). *Reducing network depth in the cascade-correlation* (Technical report No. CMU-CS-94-209). Pittsburgh: School of Computer Science, Carnegie Mellon University.

Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., & Anand, K. J. S. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: A meta-analysis. *Journal of the American Medical Association, 288*(6), 728-737.

Billiards, S. S., Pierson, C. R., Haynes, R. L., Folkerth, R. D., & Kinney, H. C. (2006). Is the late preterm infant more vulnerable to gray matter injury than the term infant? *Clinics in Perinatology, 33*(4), 915-933.

Burd, I., Chai, J., Gonzalez, J., Ofori, E., Monnerie, H., Le Roux, P. D., et al. (2009). Beyond white matter damage: Fetal neuronal injury in a mouse model of preterm birth. *American Journal of Obstetrics and Gynecology, 201*(3), 279.e1-279.e8.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven progressive matrices test. *Psychological Review, 97*(3), 404-431.

Charlton, R. A., Barrick, T. R., McIntyre, D. J., Shen, Y., O'Sullivan, M., Howe, F. A., et al. (2006). White matter damage on diffusion tensor imaging correlates with age-related cognitive decline. *Neurology, 66*(2), 217-222.

Counsell, S. J., Edwards, A. D., Chew, A. T., Anjari, M., Dyet, L. E., Srinivasan, L., et al. (2008). Specific relations between neurodevelopmental abilities and white matter microstructure in children born preterm. *Brain, 131*, 3201-3208.

Counsell, S. J., Shen, Y., Boardman, J. P., Larkman, D. J., Kapellou, O., Ward, P., et al. (2006). Axial and radial diffusivity in preterm infants who have diffuse white matter changes on magnetic resonance imaging at term-equivalent age. *Pediatrics, 117*(2), 376-386.

Dammann, O., Kuban, K. C. K., & Leviton, A. (2002). Perinatal infection, fetal inflammatory response, white matter damage, and cognitive limitations in children born preterm. *Mental Retardation and Developmental Disabilities Research Reviews, 8*(1), 46-50.

Dandurand, F., Berthiaume, V. G., & Shultz, T. R. (2007). A systematic comparison of flat and standard cascade-correlation using a student-teacher network approximation task. *Connection Science, 19*(3), 223-244.

Dyet, L. E., Kennea, N., Counsell, S. J., Maalouf, E. F., Ajayi-Obe, M., Duggan, P. J., et al. (2006). Natural history of brain lesions in extremely preterm infants studied with serial magnetic resonance imaging from birth and neurodevelopmental assessment. *Pediatrics, 118*(2), 536-548.

Eggert, T., Ladda, J., & Straube, A. (2009). Inferring the future target trajectory from visual context: Is visual background structure used for anticipatory smooth pursuit? *Experimental Brain Research, 196*(2), 205-215.

Follett, P. L., Roth, C., Follett, D., & Dammann, O. (2009). White matter damage impairs adaptive recovery more than cortical damage in an in silico model of activity-dependent plasticity. *Journal of Child Neurology, 24*(9), 1205-1211.

Heron, M., Sutton, P. D., Xu, J., Ventura, S. J., Strobino, D. M., & Guyer, B. (2009). Annual summary of vital statistics: 2007. *Pediatrics, 125*, 4-15.

Jim, K.-C., Giles, C. L., & Horne, B. G. (1996). An analysis of noise in recurrent neural networks: Convergence and Generalization. *IEEE Transactions on Neural Networks, 7*(6), 1424-1438.

LeCun, Y., Denker, J. S., & Solla, S. A. (1990). Optimal brain damage. In D. S. Touretzky (Ed.), *Advances in Neural Information Processing Systems II* (pp. 598-605). San Mateo, CA: Morgan Kaufmann.

Leviton, A., & Gressens, P. (2007). Neuronal damage accompanies perinatal white-matter damage. *Trends in Neurosciences, 30*(9), 473-478.

Leviton, A., & Paneth, N. (1990). White matter damage in preterm newborns—An epidemiologic perspective. *Early Human Development, 24*(1), 1-22.

Peterson, B. S., Vohr, B., Staib, L. H., Cannistraci, C. J., Dolberg, A., Schneider, K. C., et al. (2000). Regional brain volume abnormalities and long-term cognitive outcome in preterm infants. *Journal of the American Medical Association, 284*(15), 1939-1947.

Pierson, C., Folkerth, R., Billiards, S., Trachtenberg, F., Drinkwater, M., Volpe, J., et al. (2007). Gray matter injury associated with periventricular leukomalacia in the premature infant. *Acta Neuropathologica, 114*(6), 619-631.

Reed, R. (1993). Pruning algorithms—A survey. *IEEE Transactions on Neural Networks, 4*(5), 740-747.

Robinson, S. (2005). Systemic prenatal insults disrupt telencephalon development: Implications for potential interventions. *Epilepsy & Behavior, 7*(3), 345-363.

Schmithorst, V. J., Wilke, M., Dardzinski, B. J., & Holland, S. K. (2005). Cognitive functions correlate with white matter architecture in a normal pediatric population: A diffusion tensor MRI study. *Human Brain Mapping, 26*(2), 139-147.

Shultz, T. R. (2006). Constructive learning in the modeling of psychological development. In Y. Munakata & M. H. Johnson (Eds.), *Processes of change in brain and cognitive development: Attention and performance XXI* (pp. 61-86). Oxford: Oxford University Press.

Shultz, T. R., Mysore, S. P., & Quartz, S. R. (2007). Why let networks grow? In D. Mareschal, S. Sirois, G. Westermann & M. H. Johnson (Eds.), *Neuroconstructivism: Perspectives and prospects* (Vol. 2, pp. 65-98). Oxford: Oxford University Press.

Skranes, J., Vangberg, T. R., Kulseng, S., Indredavik, M. S., Evensen, K. A. I., Martinussen, M., et al. (2007). Clinical findings and white matter abnormalities seen on diffusion tensor imaging in adolescents with very low birth weight. *Brain, 130*(3), 654-666.

Skuy, M., Gewer, A., Osrin, Y., Khunou, D., Fridjhon, P., & Rushton, J. P. (2002). Effects of mediated learning experience on Raven's matrices scores of African and non-African university students in South Africa. *Intelligence, 30*(3), 221-232.

Unsworth, C. P., & Coghill, G. (2006). Excessive noise injection training of neural networks for markerless tracking in obscured and segmented environments. *Neural Computation, 18*(9), 2122-2145.