VARIABILITY IN THE SEVERITY OF DEVELOPMENTAL DISORDERS: A NEUROCOMPUTATIONAL ACCOUNT OF DEVELOPMENTAL REGRESSION IN AUTISM

MICHAEL SC THOMAS
VICTORIA CP KNOWLAND
ANNETTE KARMILOFF-SMITH

Developmental Neurocognition Lab, Centre for Brain and Cognitive Development, Department of Psychological Sciences, Birkbeck, University of London, Malet Street, London WC1E 7HX, UK

Developmental disorders show wide variations in severity even when, on genetic grounds, it is known that there is a common underlying cause. We use connectionist models of development combined with population modelling techniques to explore possible mechanistic causes of variations in disorder severity. Specifically, we investigate the plausibility of the hypothesis that disorder variability stems from the interaction of the common cause of the disorder with variations in neurocomputational parameters also present in the typically developing population. We base our simulations on a model of developmental regression in autism, which proposes that this phenomenon arises from over-aggressive synaptic pruning [1]. We simulated a population of 1000 networks in which 641 exhibited the behavioural marker of regression in their developmental trajectories in learning a notional cognitive domain. Aside from the known single cause of the disorder (an atypical connectivity pruning parameter), we then analysed which neurocomputational parameters contributed to variation in a number of characteristics of developmental regression. These included the timing of regression onset, its severity, its behavioural specificity, and the speed and extent of subsequent recovery. Results are related to existing causal frameworks that explain the origins of developmental deficits.

1. Introduction

Developmental disorders are notable for the range of severity with which they affect children. In the case of acquired disorders in adults, variations in the severity of behavioural deficits can usually be assigned to the degree of brain damage. For developmental disorders, the origins of variations in severity are less well understood. It is not clear, for example, whether variations in severity of deficits stem from the same causes which produce individual differences in cognitive ability in typically developing individuals. At present, many developmental disorders are defined on behavioural grounds, such as autism, dyslexia, and attention deficit hyperactivity disorder. On the face of it, the variation observed in behaviourally defined disorders could arise from at least two sources: these

disorders could just represent the bottom tail of the normal distribution in population performance on social skills or reading or attention, in which case the underlying causes of variability in the tail would be no different from those producing variation in the normal range (see, e.g., [2]); or the behaviourally defined disorders could in fact represent a heterogeneous mix of underlying causes, which would be unified by diagnosis but differ in their precise phenotype, thereby generating the observed variability in severity.

Somewhat more puzzling is the variability that is observed in developmental disorders with known genetic causes. Disorders such as Down syndrome (DS) and Williams syndrome (WS) are associated with well-characterised genetic mutations (an additional copy of chromosome 21 in the former case, a deletion of genes from one copy of chromosome 7 in the latter case). Yet, despite individuals in each disorder having a common genetic cause, they exhibit marked variation in the severity with which the genotype impacts on the cognitive profile and, more practically, on the ability of children and adults with these disorders to function in everyday life.

There are four likely sources of variation in the severity of developmental disorders of known genetic cause. (1) The environment: two individuals with the same disorder genotype might diverge through differential environmental influences. For example, phenylketonuria is a genetic disorder associated with a deficiency in an enzyme necessary to metabolize the amino acid phenylalanine to the amino acid tyrosine. This deficiency leads to a build up of phenylalanine, which causes developmental brain damage. An individual exposed to the environment of a low-phenylalanine diet will experience a much less severe version of the disorder. (2) Where genes are mutated, the genes may be polymorphic and vary in the normal population. For example, the additional chromosome 21 in DS contains many genes, some of which may differ between individuals and have different consequences. Additionally, such genetic differences may be exaggerated by gene-environment interactions. (3) The genetic mutation may occur to identical genes in two individuals, but epigenetic effects that alter gene expression might subsequently produce different effects. For example, in an extreme case, Angelman syndrome and Prader-Willi syndrome are two different disorders with the same genetic mutation (deletion of genes on chromosome 15). If the mutation is on the copy of chromosome originating from the mother, the result is Angelman syndrome; if the mutation originates from the father, the result is Prader-Willi syndrome. Parent-of-origin is an epigenetic effect that alters the gene expression from the same DNA code (in this case, the epigenetic effect is called ‘imprinting’). (4) The mutated genes are the same in two individuals, but there are individual differences in the rest of the genotype, and the mutated genes

interact with those other genes. For example, mouse models of DS have an additional copy of the equivalent of human chromosome 21 and such mice show a range of physical and cognitive abnormalities. When gene expression was examined in these mice, it was found to be altered in a large number of other genes on different chromosomes, suggesting interactions between normal and mutation genes (e.g., [3]). Once more, interactions between mutated genes and other polymorphic genes may be exaggerated by gene-environment interactions. Together, these four causes contribute to what is called the *expressivity* of a given disorder genotype, that is, the extent of the phenotypic variation given the genotype.

It should become apparent that, for the goal of explaining differences in the severity of behaviourally defined developmental disorders, the fact that we know variability occurs even when there is a common known genetic cause makes the picture more complex. That is, even if a disorder such as autism were to have a common underlying genetic cause in all individuals with autism, one would still expect individual variation in the severity of the disorder. But, of course, autism may not be a single disorder but several disorders unified by behavioural diagnostic criteria. Alternatively, in terms of underlying mechanism, autism may not be a causally distinct disorder at all, rather one end of a spectrum of normal population variation for some set of behavioural traits.

1.1. *Computational models of variability in developmental disorders*

Relatively little attention has been paid to the possible mechanistic basis of variations in the severity of developmental disorders. Thomas [4] used a connectionist pattern associator to investigate whether it was possible on behavioural grounds to distinguish between a group of simulated individuals with a common underlying disorder (plus individual differences) from a group of simulated individuals diagnosed with a disorder on behavioural grounds but actually constituting heterogeneous underlying deficits (plus individual differences). Those simulations indicated that homogeneous and heterogeneous disorder groups were not necessarily distinguishable in their mean performance levels on a range of behavioural metrics. However, they were distinguishable on the basis of group variability: the heterogeneous, behaviourally defined group had the least variance on the measure that defined the disorder while the homogeneous, common cause group had comparable variance across measures. This prediction was confirmed empirically in a comparison of naming abilities in individuals with Williams syndrome and a group of children with behaviourally defined word finding difficulties.

In another model exploring variability, Kan et al. [5] examined developmental trajectories in a small, two-node dynamical system. They demonstrated that small stochastic differences in the start state in a self-organising learning system could become exaggerated across development. From the current perspective, such stochastic differences constitute effects of the environment. Since stochastic differences occur in the brain development even of identical twins, the model explains how identical genotypes could diverge across developmental time (see [6] and [7] for similar theoretical proposals).

In this paper, we investigate the hypothesis that a common underlying genetic cause of a developmental disorder can interact with individual differences elsewhere on the genome; that these interactions can occur at the neurocomputational level; and that such interactions contribute to the variability in severity of the disorder at the behavioural level (or, indeed, whether an individual exhibits a disorder at all). This was an exploratory model with two goals: (a) to evaluate emergent effects in complex learning systems that stem from the interaction of many simultaneously interacting components, in this case serving to modulate the severity of a simulated developmental disorder; (b) to propose candidate causal mechanisms that can explain empirical observations of variations in severity, and therefore widen the range of available inferences from behaviour to causal mechanism. We took a behavioural feature found in a subset of children with autism spectrum disorder, that of *developmental regression*, and built a set of models to test our hypothesis. There were three key features of the simulations. First, there were three sources of variability in behaviour: (i) intrinsic differences in neurocomputational properties, (ii) extrinsic variability in the composition of the learning environment, and (iii) the possible presence of a disorder affecting a single neurocomputational parameter. Second, behaviour was the outcome of an extended developmental process involving interaction of the individual with a structured environment. Third, simulated individuals were classified as having a disorder on behavioural grounds.

1.2. *Developmental regression*

In developmental regression, behaviours that emerge in early development subsequently disappear. There is then a later recovery and advance of these skills to a variable level. Regression is almost unique to autism but is not a universal feature of the disorder [8]. It occurs in 20-40% of cases, with skills typically disappearing between 15 and 24 months of age, and its cause is unknown [9, 10]. Empirical data have hitherto mostly been based on retrospective parental reports, which indicate the loss of children's social and communication skills in the sec-

ond year of life, including early productive vocabulary, eye contact, gestures, reciprocal games like peek-a-boo, and sometimes a loss of play and fine motor skills.

Our simulations extended a proposal by Thomas, Knowland and Karmiloff-Smith [1] that the cause of developmental regression in autism is over-aggressive synaptic pruning. Typical brain development involves initial over-production of neural resources, and in particular, the generation of synapses which allow for the plasticity of functional circuits; and then, later in development, the pruning of unused resources [11, 12, 13]. Using a connectionist population modelling technique [14], Thomas, Knowland and Karmiloff-Smith [1] investigated the hypothesis that regression may be caused by variations in synaptic pruning. A large number of simulated individuals were exposed to an abstract learning task. One neurocomputational parameter, which determined the severity of connectivity pruning following its onset, was set to be atypically extreme. This induced developmental regression because pruning ate into functionally established pathways instead of eliminating unused connections. The extreme pruning parameter was the *sole cause of the disorder*. However, the population also incorporated background variability in a range of other neurocomputational parameters, corresponding to normal individual differences; and variability in the composition of the learning environment to which individual networks were exposed. This provided the potential for interactions between different constraints within the simulated learning systems, which might in turn modulate (negatively or positively) the severity of the regression observed in behaviour.

In the following simulations, we explored the extent to which such interactions could lead to emergent variations in (a) the timing of regression during development, (b) the severity of regression, (c) the rate at which behaviour declined, (d) its specificity to particular types of behaviours, (e) the rate at which behaviour then recovered, and (f) the final level of performance. Lastly, although our simulations assumed a genetic disorder with a single underlying cause, disorder status was behaviourally defined on the basis of observed developmental regression. We were interested in whether such sampling created a disorder population with a different correlational structure among its neurocomputational parameters compared to the full population. That is, we considered the possibility that the *sampling* inherent in behaviourally defined disorders produces 'ghost' correlations between attributes that have no direct bearing on underlying cause of the disorder (which is this case was fully known).

2. Method

2.1. *Target learning problem*

For the purposes of these simulations, the training set was considered only as an abstract mapping problem, corresponding to a notional cognitive domain. The mapping problem was quasi-regular, in that it included a predominant regularity, which could be generalised to novel input patterns, and also a set of exception patterns. The learning environment was designed to assess role of similarity, type frequency and token frequency in development, together determining the difficulty of target behaviours. The mapping problem was defined over 90 input and 100 output units, using binary coded representations. The training set comprised 508 patterns. This was complemented with a generalisation set of 410 patterns.

The predominant regularity required the network to reproduce the input pattern on the first 90 units of the output layer, and then add a binary code on the last 10 units of the output layer. There were 410 regular patterns in the training set. The regular pattern had a high type frequency and formed a consistent set of mappings, and so will be referred to as *Easy*. The generalisation ability of each network was tested on 410 novel patterns that were similar to the *Easy* patterns, in that they shared 60 of the 90 input elements. This set will be referred to as *Generalisation*. There were three different classes of exception pattern in the training set, which fell on a continuum of (dis)similarity from the predominant regularity: (1) Reproduce the input but do not add the final code [N=20]. (2) Reproduce only a portion of the input and again do not add the final code [N=68]. (3) Associate an arbitrary binary pattern with the input and again do not add the final code [N=10]. The first exception type was most similar to predominant regularity, and third type the least similar. All three possessed a lower type frequency than the predominant regularity. The combination of dissimilarity and low type frequency created a continuum of difficulty. We refer to the first exception type as *Hard*, the second as *Harder*, and the last as *Hardest*. Finally, the arbitrary mappings were sufficiently difficult that they needed to be repeated in the training set to be learned at all. They therefore provide an opportunity to assess whether greater practice provided immunity to regression or perhaps allowed better recovery from regression. The third pattern type is therefore referred to as *Hardest-practised*.

2.2. *Basic architecture*

The simulations employed a connectionist pattern associator network trained using the supervised backpropagation learning algorithm. This type of architec-

ture has been used in a number of cognitive-level models of development, for example, applied to infant categorisation, child vocabulary, semantic memory, morphosyntax, and reading development (see [15] for review). The model simulates a child's developmental profile in the notional cognitive domain.

2.3. *Variations in the learning environment*

The full training set was considered the ideal learning environment. For each individual, a subset of this training set was stochastically selected to represent the family conditions in which each simulated child was being raised. Each individual was assigned a *family quotient*, which was a number between 0 and 1. The value was used as a probability to sample from the full training set. Thus for an individual with a family quotient value of 0.75, each of the 508 training patterns had a 75% chance of being included in that individual's training set. Family quotients were sampled randomly depending on the range selected for the population. For the population we considered, the quality of the environment was reasonably good. Family quotients were sampled in the range of 0.6 to 1.0.

2.4. *Variations in the basic architecture*

Fourteen neurocomputational parameters in the basic architecture were allowed to vary between individuals, together serving to alter the learning capacity of each network. The parameter settings allowed for over 2000 billion unique individuals. Parameter values were randomly selected for each simulated individual and independently for each parameter; therefore any correlations between parameters occurred by chance. The parameters, split by their role, were as follows: *Network construction*: Architecture (two-layer network, three-layer network, or a fully connected network incorporating a layer of hidden units plus direct input-output connections); number of hidden units (10 to 500); range for initial connection weight randomisation (±0.01 to ±3.00); sparseness of initial connectivity between layers (50% to 100% of potential connectivity). *Network activation*: unit threshold function (logistic function, sigmoid temperatures between 0.0625 and 4); processing noise (0 to 6); response accuracy threshold (0.0025 to 0.5). *Network adaptation*: backpropagation error metric (Euclidean distance or cross-entropy); learning rate (0.005 to 0.5); momentum (0 to 0.75). *Network maintenance*: weight decay (0 to 2×10^{-5} per pattern presentation); pruning onset (0 to 1000 epochs); pruning probability (0 to 1); pruning threshold (0.1 to 4.0). Detailed explanations of the role of each parameter can be found in [14].

The pruning process was central to simulating developmental regression. Networks were created with initial connectivity, determined by the sparseness

parameter and the weight variance parameter. After a number of epochs of training determined by the pruning onset parameter, pruning commenced. After each epoch of training following onset, all the connection weights were assessed. Any connection whose magnitude was less than the pruning threshold was deemed an unused resource and could be permanently pruned. If a connection was less than threshold, pruning was then stochastic, occurring with a probability determined by the pruning probability parameter. Benchmarking suggested that of the 14 parameters varying in the simulations, the pruning threshold was the sole parameter that produced developmental regression. At levels up to 1.0, little regression was found (for comparison, initial weights were most often randomised in the range ±0.5). Levels above 1.0 were associated with regression. In the following population, pruning thresholds varied up to 4.0. This range implemented the hypothesis that the cause of the disorder is an accumulation of risk gene variants that allows a neurocomputational parameter to take on more extreme values than found in unaffected individuals. Under the hypothesis, both genetic risk and the neurocomputational causal factor are continuously valued.

2.5. *Creation of a population*

Parameter sets for 1000 individuals were generated at random. A family quotient value was generated in the appropriate range and the quotient used to create each individual's bespoke family training set. Each network was initiated with random weight values (in the range determined by the individual's weight range parameter), and then trained for 1000 epochs, where one epoch was an exposure to all the patterns in the individual's training set, presented in random order. Performance was measured on the five pattern types (*Easy, Generalisation*, *Hard*, *Harder*, *Hardest-practised*) according to the full training set and the full generalisation set.

3. Results

Developmental regression was defined on behavioural grounds as a noticeable drop in performance over development in one or more of the five behavioural measures (*Easy*, *Generalisation*, *Hard*, *Harder*, and *Hardest-practised*) against the level of variability exhibited by a given simulated individual. Developmental trajectories were plotted for all 1000 individuals and coded by hand for whether or not they exhibited developmental regression. Hand coding was used because trajectories were often non-monotonic and noisy, rendering automated classification problematic. Double coding was carried out on a sub-sample to ensure consistency. Figure 1 depicts sample trajectories for four typically developing net-

works (defined here as those not exhibiting regression) and four networks that showed regression. When regression was observed for a given pattern type, six measurements were made: peak performance prior to regression; the epoch at which regression occurred; the size of the drop in accuracy; the number of epochs over which that drop occurred; the rate of recovery (five qualitative categories were used: no recovery, slow recovery, medium recovery, fast recovery, almost instant recovery); and the final level of performance at the end of training (1000 epochs). We defined *four levels of severity* of the disorder, based on the size of the decline in accuracy: level 1, corresponding to a drop in accuracy of between 0 and 20%; level 2, corresponding to a drop in accuracy of 20 to 40%; level 3, corresponding to a drop between 40 and 60%; and level 4, corresponding to a drop of 60 to 100% in accuracy. Of the 1000 simulated individuals, 641 cases of regression were recorded in one or more behaviour.

3.1. *Variations in severity*

The single cause of regression was increasing the size of the pruning threshold parameter. However, there was no direct relationship between increasing this parameter and the subsequent severity of the observed regression. Table 1 lists the probability of observing regression at each severity level, for the five separate mapping types. It is evident that increases in the value of the pruning threshold produce an increased liability to exhibit regression, and regression of a more severe level, but the outcome was not deterministic (e.g., not 100% likely even with the highest threshold). Comparison of the behavioural categories indicated that *Easy* and *Generalisation* were more robust to regression, with all hard patterns more vulnerable. Notably, the *Hardest-practised* mapping exhibited a bimodal distribution for most extreme pruning levels: regression was either absent (23%) or at the severest level (61%). This is because the combination of exception mappings and extra practice caused networks to learn such mappings using a small number of weights (more localist representations). These weights were then stochastically preserved or lost during the atypical pruning process. More distributed representations demonstrated a dose-response relationship.

A statistical stepwise logistic regression analysis was used to identify the relationship between variation in neurocomputational parameters and the *presence of regression*. As expected, the atypical setting of the pruning threshold parameter accounted for most of the variance (Nagelkerke R^2 = 59.7% of the variance explained).

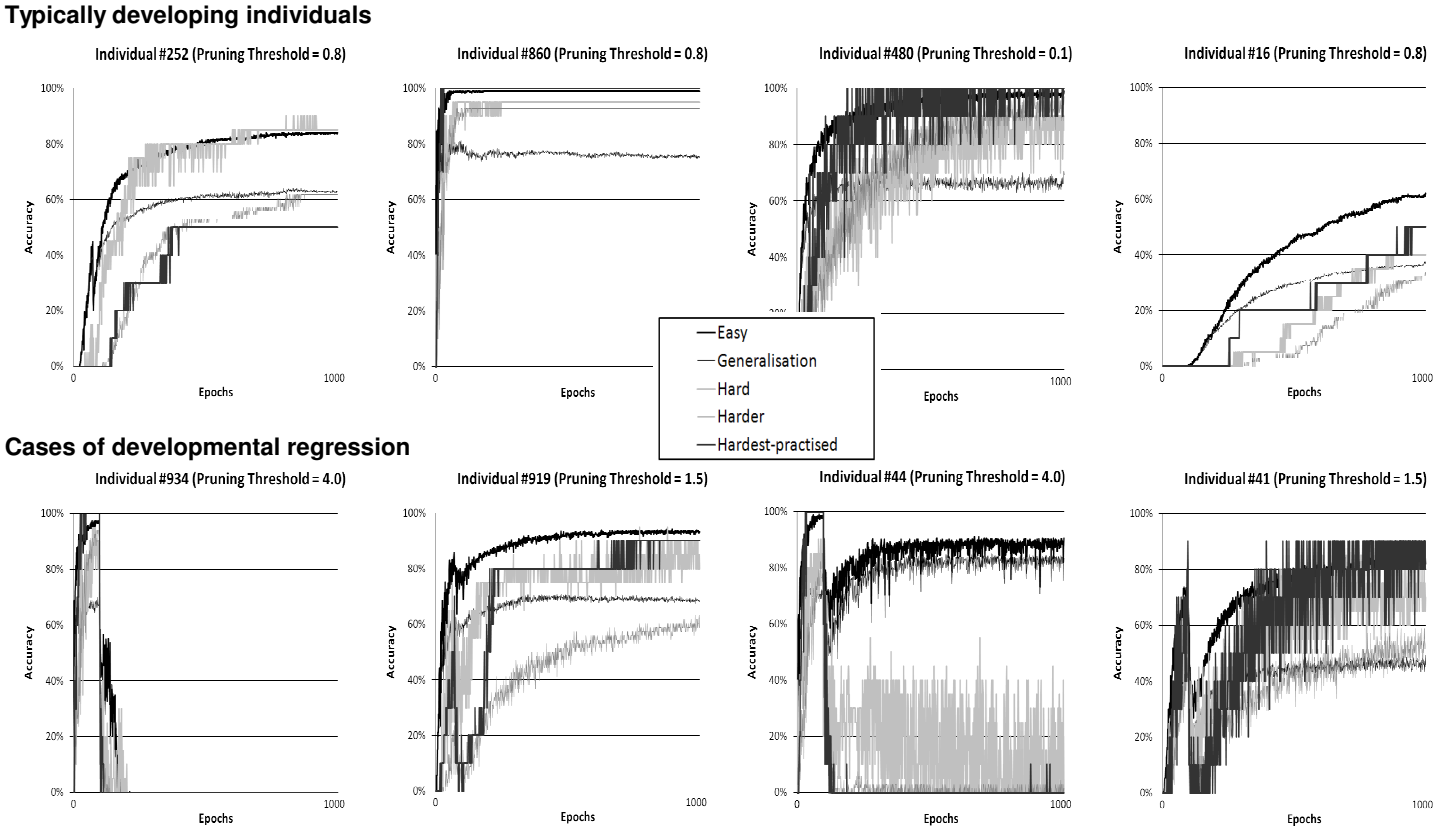

Figure 1. Illustrative developmental trajectories for four typically developing networks and four networks exhibiting regression.

Table 1. The probability of a network exhibiting developmental regression, split by the level of severity (None, least severe I, II, III, or most severe IV). Probabilities are shown separately for each pattern type.

		Pruning Threshold								
Severity	Pattern	.1	.75	1.0	1.5	2.0	2.5	3.0	3.5	4.0
None	*Easy*	.98	.90	.72	.35	.24	.16	.14	.14	.04
	Generalisation	.98	.90	.72	.40	.33	.22	.17	.17	.10
	Hard	.98	.95	.75	.37	.29	.23	.16	.22	.11
	Harder	.98	.90	.71	.34	.23	.19	.13	.17	.11
	Hardest-Practised	.98	.95	.80	.42	.31	.19	.27	.23	.23
Severity I	*Easy*	.00	.07	.22	.24	.20	.17	.09	.09	.08
	Generalisation	.01	.03	.03	.24	.27	.20	.18	.22	.11
	Hard	.01	.00	.02	.11	.16	.20	.20	.15	.25
	Harder	.00	.00	.02	.06	.12	.28	.39	.40	.52
	Hardest-Practised	.01	.09	.24	.30	.20	.22	.13	.13	.04
Severity II	*Easy*	.01	.01	.02	.20	.30	.25	.28	.26	.24
	Generalisation	.00	.00	.01	.08	.12	.18	.20	.22	.25
	Hard	.00	.00	.02	.02	.05	.13	.23	.22	.36
	Harder	.00	.00	.02	.02	.01	.01	.03	.03	.05
	Hardest-Practised	.00	.04	.11	.10	.09	.07	.12	.05	.05
Severity III	*Easy*	.02	.01	.10	.24	.27	.28	.21	.17	.18
	Generalisation	.00	.00	.02	.28	.34	.41	.49	.54	.61
	Hard	.00	.00	.04	.06	.08	.08	.17	.16	.18
	Harder	.01	.08	.16	.19	.13	.24	.14	.15	.10
	Hardest-Practised	.00	.00	.06	.24	.30	.23	.17	.13	.13
Severity IV	*Easy*	.01	.02	.03	.16	.26	.26	.38	.40	.48
	Generalisation	.00	.01	.05	.05	.02	.04	.06	.05	.08
	Hard	.00	.01	.09	.06	.08	.07	.05	.07	.04
	Harder	.02	.02	.02	.18	.13	.18	.10	.12	.04
	Hardest-Practised	.00	.02	.05	.28	.46	.52	.51	.53	.61

However, individual differences in other neurocomputational parameters also modulated the risk. The following parameters showed a significant relationship: unit threshold function (Nagelkerke R^2 = 6.8%), architecture (1.2%), pruning probability (1.0%), hidden unit number (0.5%), sparseness (0.3%) and momentum (0.3%). Variations in the family quotient did not reliably modulate regression risk. In the case of pruning probability, this parameter directly affected the pruning process. All the other parameters acted indirectly by increasing or decreasing the size of the connection weights produced by learning prior to the onset of pruning. A statistical stepwise linear regression analysis was used to identify which parameters predicted the *severity of regression*, that is, the size of

the drop in accuracy during regression. Twelve computational parameters were significant contributors at the .05 level; only 6 individually accounted for more than 1% of the variance: pruning threshold (10.0%), unit threshold (5.8%), processing noise (2.5%), architecture (2.3%), pruning probability (1.9%) and learning algorithm (1.4%). Pattern type explained 3.9% of the variance and family quotient 0.8%. Once again, the value of the atypical pruning parameter explained much of the variance in severity of regression, but many other parameters varying in the population as a whole also contributed to predicting severity.

3.2. *Variability in onset*

The onset of regression approximated a normal distribution with a mean of 101 epochs and a standard deviation of 69.[a] A statistical linear regression analysis indicated that the pruning onset parameter explained 51.1% of the variance. Only unit threshold and learning rate (0.3% each) also reached significance. Regression, then, was triggered by the normal onset of the pruning process.

3.3. *Variability in the rate of decline*

The rate of decline was assessed for each pattern type, separately across severity levels. More severe regression showed slower declines, and the *Hardest-practised* showed fastest decline amongst the pattern types. The statistical stepwise linear regression analysis implicated 8 computational parameters in explaining the variability in how quickly performance declined, with only 3 explaining more than 1% of the variance: pruning probability (5.3%), pruning threshold (1.5%) and unit threshold (1.5%). The amount of unexplained variance in this case reflected the stochastic nature of the pruning process.

3.4. *Variability in the rate of recovery*

Recovery rates were faster for milder regression and for *Easy* and *Generalisation* patterns. High type frequency and consistent mappings were thus recovered more easily. The statistical stepwise linear regression analysis implicated 8 computational parameters in explaining the variability in recovery rates, with 4 accounting for more than 1% of the variance: pruning threshold (9.9%), unit threshold (2.2%), processing noise (1.7%) and response accuracy threshold (1.4%). Pattern type explained 8.8% of the variance and family quotient 0.2%. Recovery, then, depended on how severe the loss of resources had been, and the

[a] Plots illustrating population variation for data reported in Sections 3.2 to 3.6 can be found at http://www.psyc.bbk.ac.uk/research/DNL/techreport/NCPW12plots.pdf

difficulty of the behaviour being recovered. It was little affected by differences in the quality of the environment, indexed by the family quotient parameter.

3.5. *Variability in recovery levels*

The level of recovery can be considered in two ways: either the final outcome level or how far the network has recovered behaviours compared to their pre-regression peak. The former includes the differential difficulty of learning each pattern type while the latter corrects for this difference. The statistical stepwise linear regression analysis indicated that up to 13 of the 14 computational parameters were implicated in explaining the variability in recovery levels. The parameters split into three groups: those relating to the *level of damage*, those relating to the *level of background plasticity*, and those relating to *quality of processing*. Focusing on the relative measure, the main contributors were pruning threshold (21.6%) and pruning probability (0.9%), both indexing damage; unit threshold (10.7%) indexing plasticity; and processing noise (3.7%) indexing quality of processing, respectively. Pattern type explained 6.2% of the variance, with *Easy* and *Generalisation* showing higher recovered levels. Notably, practise for the *Hardest* patterns did not benefit recovery. Family quotient explained only 0.2% of the variance in recovery.

3.6. *Sampling and correlational structure*

In the full population, the 14 parameters that defined each individual network were sampled independently and at random (see Methods). We did not, therefore, expect any systematic relationship between parameters in the full population, other than those spuriously generated by multiple comparisons. Table 2 shows that three correlations were significant at the .01 level in the full population. We then repeated the correlations on just those 641 individuals who exhibited the behavioural marker of regression. The former three correlations were no longer reliable at the .01 level, but five new reliable correlations appeared. Two of these involved the pruning threshold parameter, which was the cause of regression. Three correlations were between parameters involved in background variation, which when interacting with each other could together elevate risk for regression. For example, sparse connectivity was a protective factor because it encouraged larger connection sizes more resistant to pruning, but not if the learning rate was low. Notably, then, the simulations suggest that the *selective sampling* involved in a behaviourally defined disorder can create correlations between processing parameters that are related, but sometimes only indirectly, to the underlying cause of the disorder.

Table 2. Correlations between neurocomputational parameters either in the full population of 1000 networks, just those exhibiting developmental regression (N=641), or those not exhibiting regression (N=359). Given the large number of potential correlations, only those with p<.01 (2-tailed) in either the full or sub-populations are included. Manipulation to the pruning threshold parameter was the sole cause of regression. Values show Pearson Correlation and significance in brackets.

Parameter 1	Parameter 2	Whole population (N=1000)	Those exhibiting Regression (N=641)	Those not exhibiting Regression (N=359)
Pruning threshold	Unit threshold	.046 (.143)	-.166 (<.001)*	-.142 (.007)*
Pruning threshold	Momentum	.031 (.325)	.118 (.003)*	.046 (.389)
Response accuracy	Weight variance	.091 (.004)*	.090 (.023)	.092 (.083)
Pruning onset	Learning rate	-.021 (.515.)	-.105 (.008)*	.095 (.073)
Pruning onset	Weight variance	.087 (.006)*	.086 (.030)	.088 (.097)
Pruning onset	Learning algorithm	-.043 (.171)	-.120 (.002)	.052 (.329)
Pruning onset	Sparseness	.084 (.008)*	.100 (.011)	.064 (.229)
Learning rate	Sparseness	-.063 (.046)	-.104 (.008)	.009 (.872)

* significant at .01 level (2-tailed)

4. Discussion

In the model, the disorder of developmental regression was caused by a single underlying computational cause, over-aggressive pruning of connectivity [1]. Nevertheless, background variability in other neurocomputational parameters present in the whole population led to marked variation in the severity and characteristics of the disorder, such that there was no deterministic relationship between the primary underlying cause and its surface manifestation in behaviour (see [16] for related work in the field of resilience to cognitive decline in ageing). Statistical analyses revealed that the relevant properties of background neurocomputational variation depended on the behavioural characteristic under consideration. Variance in some behavioural characteristics was predicted by only a few parameters: the onset of pruning and the speed of the decline involved only the details of the pruning mechanism. Variance in other behavioural characteristics, such as the severity of regression and the speed and final level of recovery, involved contributions from a wide range of parameters, including those involved in network plasticity and the quality of processing. Moreover, the type of behaviour also influenced the characteristics of regression. Mappings with higher type frequency and consistency were more robust and, if lost, faster to recovery. Idiosyncratic mappings were hard to recover even with practice. Differences in the composition of the learning environment, at least for the range considered, explained little of the variance in regression or recovery. Lastly, we

noted that even when there was a single underlying mechanistic cause, diagnosis of a disorder on behavioural grounds created a ghost correlational structure between the neurocomputational properties of affected individuals that were an artefact of sampling, and only indirectly related to the underlying cause.

In the Introduction, we identified four possible sources of variability in the severity of disorders with a single underlying cause. The current simulations showed some evidence for environmental effects (source #1), albeit not from variations in the learning environment but from stochastic events (e.g., *Hardest-practised* patterns showed either no or very severe regression depending on whether by chance certain key connections were pruned). We presented extensive evidence that interactions with background population-wide variability could confer a probabilistic relationship between underlying cause and behavioural manifestation (source #4). However, in the simulations, interactions occurred at a neurocomputational rather than genetic level. Gene-environment interactions emerged in training pattern effects, which modulated both regression and recovery. However, statistical analyses revealed little evidence that variations in the composition of the learning environment (implemented here by family quotient) contributed to the severity of the disorder (at least, when the learning environment was on the whole of good quality; see [1] for a consideration of the effects of very impoverished learning environments). Since the underlying cause of regression was taken to be an accumulation of risk genes present in the normal population rather than a genetic mutation, we did not model the possibility of disorder variability arising from variations in mutated genes (source #2). Epigenetic effects were also beyond the scope of the simulations (source #3).

More widely, the causal account of disorders we have offered is consistent with recent proposals by Bishop [17] that the medical model is more appropriate than adult neuropsychological dissociation methodology for understanding the origin of developmental deficits. In the medical model, disease status is conveyed probabilistically based on risk and protective factors, rather than necessary and sufficient causes.

The value of the current model is threefold. First, it demonstrates a mechanistic rather than statistical basis for why the relationship between disorder cause and behavioural outcome should be non-deterministic. Second, through implementation, it provides a framework to investigate this mechanistic basis in quantitative rather than qualitative terms. The investigation revealed, for example, that separate individual difference factors might contribute to variations in regression onset and speed, compared to severity and recovery. Third, the simulations showed the benefit of considering disorders within a population setting, where variability is a dependent variable rather than a source of noise to be over-

come by averaging, and where individual differences in symptom severity are a target phenomenon to be explained. This approach sets the stage for using computational models of development to predict interventions that are tailored to individual manifestations of a disorder rather than to an idealised average disorder that may not be observed in any individual child.

Acknowledgments

This research was supported by ESRC grant RES-062-23-2721 and MRC grant G0300188 to the first author and an ESRC studentship to the second author.

References

1. M. S. C. Thomas, V. Knowland & A. Karmiloff-Smith, *Manuscript submitted for publication* (2010).
2. Y. Kovas, C. Haworth, P. Dale & R. Plomin, *Monographs of the SRCD*, **72(3)**, serial. No. 188, 1-144 (2007).
3. N. Saran, M. Pletcher, J. Natale, Y. Cheng & R. Reeves, *Human Molecular Genetics*, **12(16)**, 2013-2019 (2003).
4. M. S. C. Thomas, *Dev. Sci.*, **6(5)**, 537-556 (2003).
5. K. Kan, A. Ploeger, M. Raijmakers, C. Dolan & H. van der Maas, *Dev. Sci.*, **13**, 11-27 (2010).
6. A. Karmiloff-Smith, *Trends in Cog. Sci.*, **2**, 389-398 (1998).
7. A. Oliver, M. Johnson, A. Karmiloff-Smith & B. Pennington. *Dev. Sci*., **3**, 1–23 (2000).
8. A. Pickles, E. Simonoff, G. Conti-Ramsden, M. Falcaro, Z. Simkin, T. Charman, et al., *JCPP*, **50**, 843-852 (2009).
9. G. Baird, T. Charman, A. Pickles, S. Chandler, T. Loucas, et al., *J. Autism Dev. Disorder*, **38**, 1827-1836 (2008).
10. J. Richler, R. Luyster, S. Risi, W. Hsu, G. Dawson, et al. (2006). *J. Autism and Dev. Disorders*, **36(3)**, 299-316. (2006).
11. M. Johnson, *Developmental cognitive neuroscience (2nd Ed.)*. Oxford: Wiley-Blackwell (2005).
12. P. Huttenlocher, *Neural plasticity*, Cam. Mass.: Harvard Uni. Press (2002).
13. P. Huttenlocher & A. Dabholkar, *J. Comp. Neurol*., **387**, 167–178 (1997).
14. M. S. C. Thomas, A. Ronald & N. Forrester, *Manuscript submitted for publication* (2011).
15. D. Mareschal & M. S. C. Thomas, *IEEE Transactions on Evolutionary Computation*, **11(2)**, 137-150 (2007).
16. M. S. C. Thomas, *Proc. 30th Annual Conference of the Cog. Sci. Soc*., 2089-2094 (2008).
17. D. Bishop, *Qu. J. Exp. Psych*., **59**, 1153-1168 (2006)