Additional Plots and Subsidiary analyses for

Thomas, M. S. C., Knowland, V. C. P., & Karmiloff-Smith, A. (2011). Variability in the severity of developmental disorders: A neurocomputational account of developmental regression in autism. In E. Davelaar (Ed.), *Proceedings of the 12th Neural Computation and Psychology Workshop* (p. x-y). World Scientific.

Original paper abstract:

Developmental disorders show wide variations in severity even when, on genetic grounds, it is known that there is a common underlying cause. We used connectionist models of development combined with population modelling techniques to explore possible mechanistic causes of variations in disorder severity. Specifically, we investigated 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 wider 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 (Thomas, Knowland & Karmiloff-Smith, 2011). We simulated a population of 1000 networks in which 641 exhibited the behavioural marker of regression in their developmental trajectories when 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 observed 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 to explain the origins of developmental deficits.

Additional material:

The additional figures included in the current document depict the variability observed in characteristics of the simulated developmental regression, included the timing of its onset (Figure A), the time period over which regression took place indexing the rate of decline (Figure B), the rate of subsequent recovery split by the severity of initial regression (Figure C), and in the final level of performance either in absolute terms or relative to the level of performance achieved prior to the onset of regression (Figure D).

The subsidiary analyses exploit the fact that simulated individuals had *sibling relationships*. Artificial genomes were used to encode each individual's inherited neurocomputational parameter set. Siblings shared 50% similarity in their genomes, and therefore had (probabilitistically) similar parameter sets. Where an individual demonstrating regression had an unaffected sibling, it was then possible to explore whether the learning ability of the unaffected sibling accounted for any variability in the severity of the regression found in the

affected individual. The subsidiary analyses addressed two questions: (1) Does unaffected sibling ability explain differences in the severity of the disorder in the affected individual? (2) Does unaffected sibling ability serve as a protective factor in whether the cause of the disorder (a high setting of the pruning threshold parameter) actually led to manifestation of the disorder?

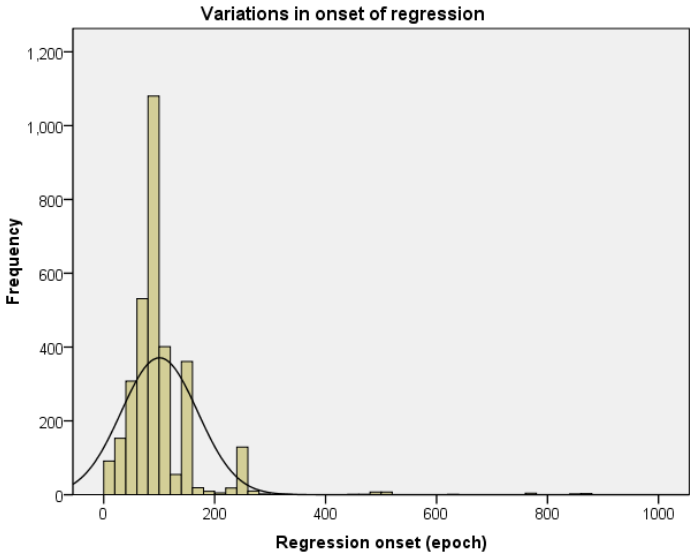
The results were as follows: (1) Unaffected sibling ability was defined as the sibling's rank order in the unaffected population on one of the harder mapping sets at an early point in training (50 epochs). This measure served to discriminate learning ability across individuals. Sibling rank order was a weak predictor of disorder severity, accounting for 8% of the variance. The simulations suggest *some degree of the severity of an individual's disorder can be explained by their unaffected sibling's cognitive ability*. (2) Unaffected sibling rank order did not serve to modulate the relationship between the parameter causing regression and whether an individual actually exhibited regression. However, for unaffected individuals only, there was a reliable relationship between the value of this parameter and their unaffected sibling's ability. Somewhat surprisingly, a lower ranking in the unaffected sibling correlated with a value of the parameter that placed the individual more at risk of regression. Our interpretation was that in these simulations, *a genetic family background conferring LOW learning ability was a protective factor again regression*. Should regression index a mechanism responsible for the wider autistic phenotype, *the sibling results suggest that autism would be associated with high intelligence families rather than low intelligence families*.

Additional Plots

Additional plots for Sections 3.2-3.5

Section 3.2

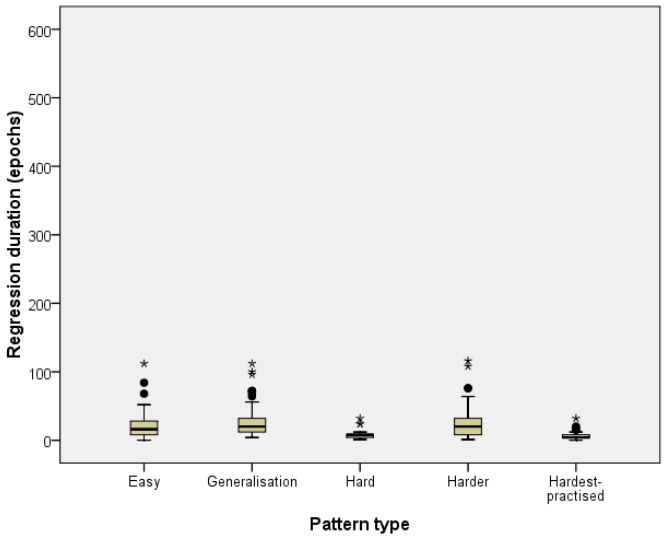
Figure A. Distribution of onsets of developmental regression (the 'life' of each individual network was 1000 epochs of training). For each box, the central line represents the median value of the group; the box captures the middle 50% of the cases; the whiskers connect the largest and smallest values that are not categorized as outliers or extreme values; 'o' represents an outlier more than 1.5 box-lengths away from the box; '*' represents an extreme value more than 3 box-lengths away from the box.



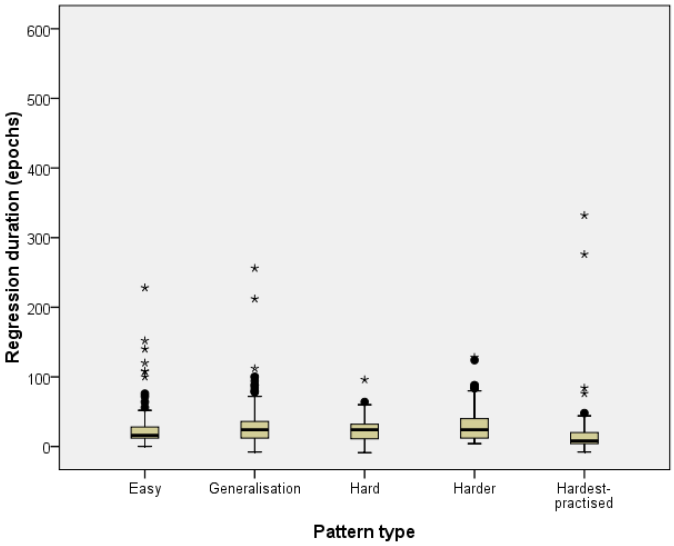
Section 3.3

Figure B. Variation of the number of epochs over which the regressive decline in performance took place, split by severity and shown for each pattern type.

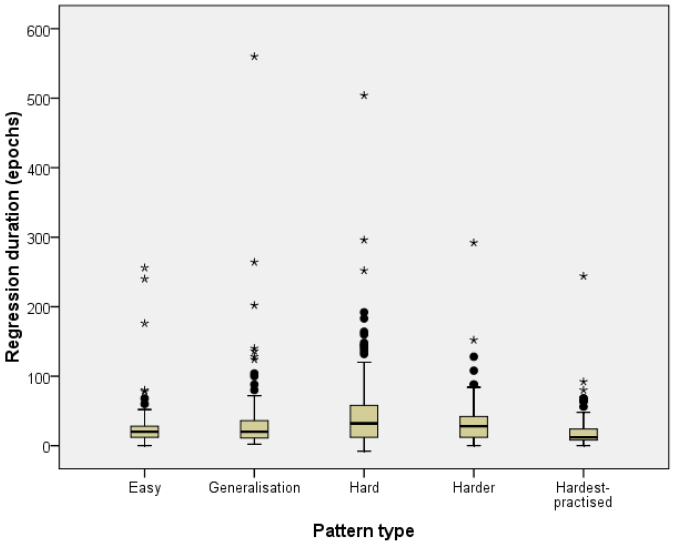
Severity 1



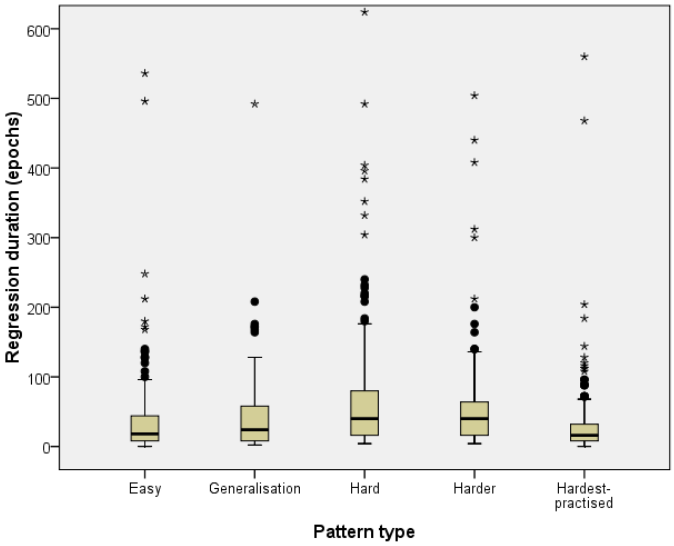
Severity 2



Severity 3



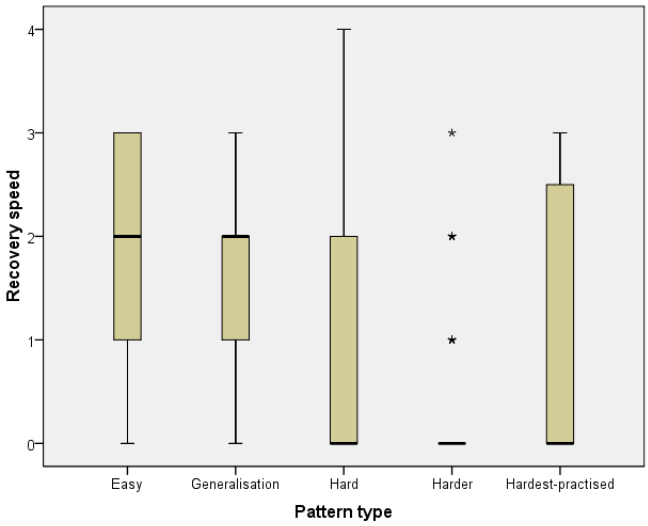
Severity 4



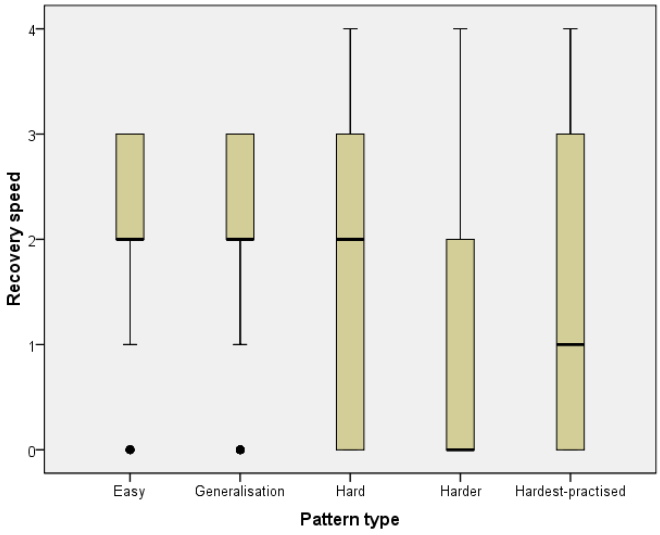
Section 3.4

Figure C. Variation in the rate of recovery from regression, split by severity and shown for each pattern type. Recovery was categorised at five levels: 0 = no recovery, 1 = slow recovery, 2 = medium recovery, 3 = fast recovery, 4 = very fast recovery.

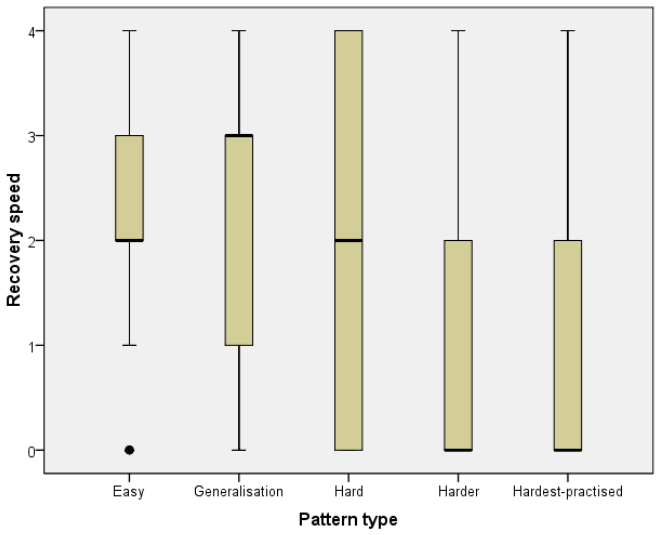
Severity 1



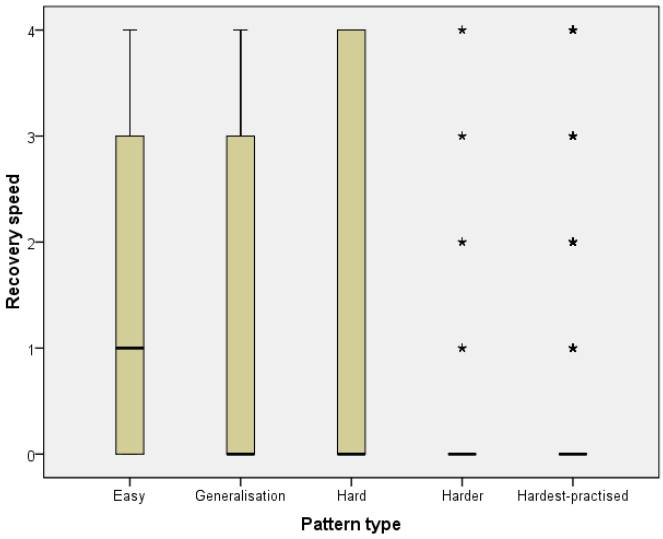
Severity 2



Severity 3



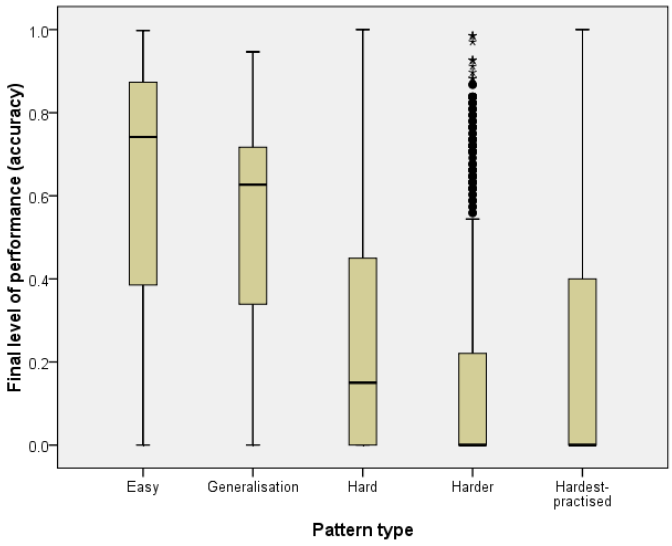
Severity 4



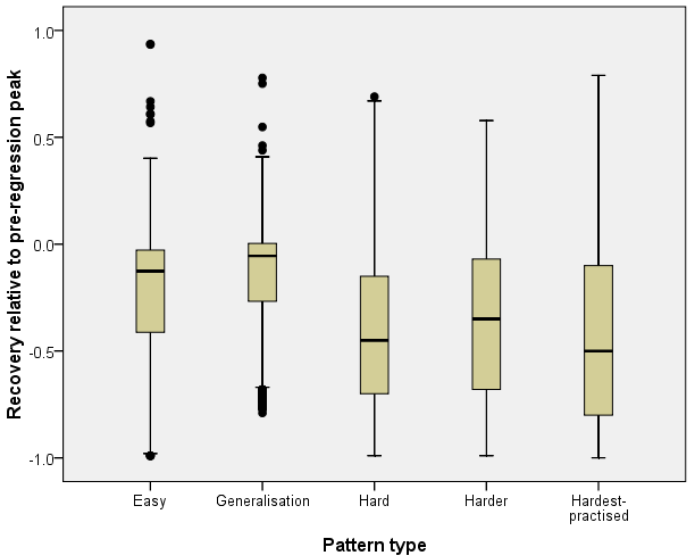
Section 3.5

Figure D. Variation in the recovery levels following regression. (a) Absolute final level of performance; (b) Level of final performance relative to the pre-regression peak (0 = return to the peak level; negative values = permanent deficit; positive values = later development above peak value).

(a) Absolute recovery



(b) Relative recovery



Subsidiary analyses

Additional analysis evaluating whether differences in disorder symptom severity can be predicted from the ability of unaffected siblings.

If a sibling is more able, might this family background convey resilience to individuals affected by the disorder? An artificial genome was used to encode each individual's neurocomputational parameter set. This allowed for the simulation of genetically similar individuals such as siblings. We used this manipulation to explore whether the 'ability level' of unaffected siblings of networks showing regression could be used to predict the level of severity of regression in the affected individual. If a sibling is more able, might this family background convey resilience to individuals affected by the disorder, so that the disorder was milder? This might point to a method to assess the contribution of normal individual differences (i.e., background genetic variability) to disorder severity.

Additional Methods

In order to allow *siblings* to be simulated, parameter values were encoded in an artificial genome. Siblings were defined by their genetic similarity. Each parameter was encoded in a set of binary genes, with the number of 1-valued alleles from the set determining the parameter value via a look-up table. For example, hidden unit number was coded over ten binary genes. If an individual had a genotype of 0110101100, a total of five 1s corresponded to a hidden layer with 60 units. A look-up table was created for each parameter (available in Thomas, Ronald & Forrester, 2011). Sibling pairs then constituted genomes that shared 50% of their genes, constraining the neurocomputational parameters to be similar.

The full population was generated in the following manner. Five hundred genotypes were generated at random. From these, five hundred further sibling pairs were generated. The look-up tables were used to produce the parameter set for each individual. A family quotient value was generated in the appropriate range for that population and the quotient was then used to create each individual's bespoke family training set. Siblings were exposed to the same 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 training set, presented in random order. Performance was measured on the five pattern types (*Eas*y, *Generalisation*, *Hard*, *Harder*, *Hardest-practised*) according to the full training set and the generalisation set.

Variations in the learning environment were implemented via the family quotient parameter, a value created for each family that was used to produce a probabilistically sampled subset of the full training set (e.g., a family quotient value of 0.8 would be used to select 80% of the full training set). Family quotients were sampled in the range of 0.6 to 1.0. Siblings were assumed to be raised in the same family and were assigned the same training set.

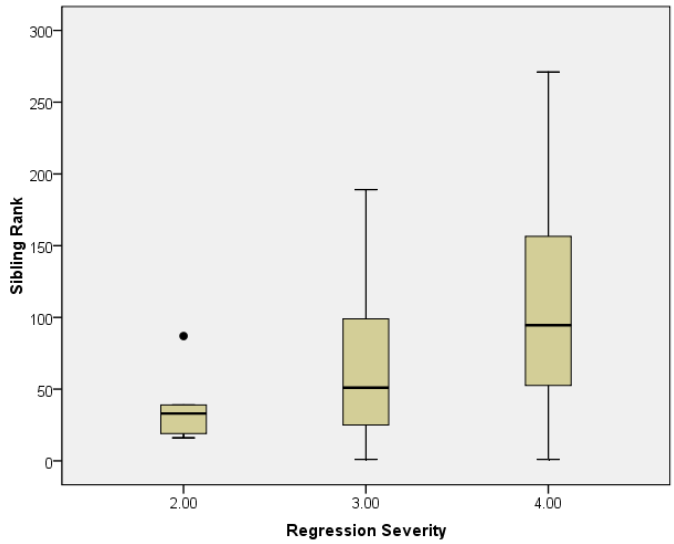
Results

Unaffected siblings as predictors of disorder symptom severity

We took a point in early development (50 epochs) and one of the more discriminating behaviours (the *Harder* patterns) and gave the typically developing networks a rank order in the population. This gave a proxy of the unaffected siblings 'cognitive ability'. We then identified sibling pairs for whom one individual showed regression (of any level in any behaviour), while the other showed no regression of any severity in any behaviour. The possible reasons why a sibling could be unaffected were that background protective factors had conspired to save the sibling despite a high pruning threshold, or that the pruning threshold was not as high in the sibling, or a combination of both.

We then tested whether the rank order of the unaffected sibling in the population (their ability level) predicted the severity of the regression in the affected sibling. We excluded the more potentially ambiguous condition of mild regression. For severity levels 2, 3, and 4, the numbers of affected siblings at each severity level were unequal, compromising the analysis (N=5, 12, and 52, respectively). A statistical linear regression analysis was used to predict severity from rank order. One of the 69 data points had a Cook's distance of .4, while the other 68 had values below .1, suggesting that this data point was an outlier. Removing the outlier, population rank order predicted 8.1% of the variability in severity of regression ($F(1,67)=5.92$, $p=.018$). This is either an encouraging or discouraging result. Encouragingly, unaffected sibling ability, as a measure of background genetic individual variation, reliably predicted variability in disorder severity. However, over 90% of the variance was left unexplained.

Figure E: The population rank at 50 epochs on the *Harder* patterns for unaffected 'siblings' of networks showing regression, split by level of severity of regression in the affected sibling.



Unaffected sibling ability as a direct modulator of regression risk

We next asked whether unaffected sibling ability modulated the risk of the process causing regression actually leading to a manifestation of the disorder. We know that the cause of regression in affected networks was an atypical setting of the pruning threshold parameter. If the value is high enough, connections in functionally established circuits become liable to pruning, rather than unused resources. In previous analyses, we found that the value of the pruning threshold parameter explained around 60% of the risk in the manifestation of regression. We therefore evaluated whether unaffected sibling ability modulated the relationship between the value of the pruning threshold parameter and whether an individual exhibited regression (in any of the five target behaviours, at any level of severity).

For cases where one sibling demonstrated regression but the other was unaffected, a linear regression model was used to predict regression status (yes or no) using the predictor of pruning threshold, the predictor of unaffected sibling ability, and an interaction term of threshold*sibling ability. The results indicated that as expected, pruning threshold was a reliable predictor of regression ($F(1, 193)=82.98$, $p<.001$, partial-eta squared=.301). However, neither sibling ability ($F(1, 193)=.09$, $p=.763$, partial-eta squared=.000) nor the interaction term explained a significant amount of regression risk ($F(1, 193)=.27$, $p=.604$, partial-eta squared=.001). Unaffected sibling ability did not, therefore, modulate the relationship between the parameter causing regression and the risk of regression occurring.

Sibling ability as a protective factor operating in unaffected individuals

Lastly, we considered whether family background might be serving as a protective factor: perhaps our analysis was failing to consider those individuals who were at risk of experiencing regression, but did not because their family background protected them from it. (In affected individuals, the protective factors were not sufficient).

We considered the correlation between unaffected sibling ability and the pruning threshold just in those simulated individuals who developed normally. Our initial expectation was that if family genetic background, as measured via unaffected sibling ability, protected individuals again regression, then we might find higher values of the pruning threshold for those with more able siblings. (Without a good family genetic background, regression would have occurred, and the individual would not be present in this unaffected group). We predicted a negative correlation between sibling rank and pruning threshold in the unaffected group (where low rank = good performance and high pruning threshold = risk for regression).

There was indeed a reliable relationship between unaffected sibling ability and pruning threshold in typically developing individuals, explaining 3% of the variance ($N=197$, $R^2=.030$, $F(1,195)=6.07$, $p=.015$). However, surprisingly, this was a positive correlation

(of .174). That is, higher values of the pruning threshold parameter were present in typically developing individuals with poorer performing siblings. No such relationship between pruning threshold and unaffected sibling ability was found in those individuals who did display regression (N=303, p>.4). Since siblings shared the family environment, it is possible that the family protective factor was carried by the family environment rather than the genetic background. While it was indeed that case that the quality of the environment explained some of the unaffected sibling's ability (the Pearson correlation between family quotient and sibling rank was .205, p=.004), nevertheless the quality of the environment did not correlate with the pruning threshold parameter in the unaffected individuals (Pearson correlation = -.122, p=.088) and therefore could not explain the predictive power of sibling ability.

Discussion

(1) We demonstrated that if background variability in the neurocomputational properties of a learning system is determined by genotype, then the abilities of unaffected siblings could explain some degree of the variation in the *severity* of the regression found in the affected sibling, although over 90% of the variance was unexplained by this method.

(2) We found that unaffected sibling ability did not directly modulate the relationship between the causal process responsible for producing regression in the population (aggressive connectivity pruning) and the actual incidence of regression.

(3) However, we did find that in those *unaffected* by regression, sibling ability was associated with higher values of the relevant pruning parameter, as if ability were successfully serving as a protective factor against regression. Against initial expectations, the relationship was such that a genetic family background conveying *low ability* was associated with values of the pruning parameter conveying greater risk of regression.

There are two interpretations of this unexpected result. Either: a high pruning threshold simply slowed learning; unaffected siblings were both at risk of having higher pruning thresholds; hence one sibling's high threshold was associated with the other's slow learning / low ability. If this were correct, it is puzzling that the relationship was only found in unaffected individuals, not in affected individuals as well. Under this view, affected individuals should only differ because their pruning threshold happened to lead to regression as well; why should their sibling with a (genetically similar) threshold also not then show slow learning.

The alternative interpretation is that *low ability conveyed a protection against regression*, so that higher pruning threshold values were possible in the unaffected group when other inherited parameters produced low ability (using the sibling's ability as an index of that ability). In these individuals, protection outweighed risk. This second interpretation is feasible because slow learning networks are often low capacity networks. Such networks must develop larger weights to learn the domain to which they are exposed. This is because the fewer connections must become proportionally larger to deliver the

equivalent input-output mappings. Large connection weights confer protection against a pruning process that removes small weights.

If this second interpretation is correct and, as we have argued, regression actually indexes a mechanism responsible for the wider autistic phenotype, then the sibling simulations make the following broader prediction: *autism should be associated with high intelligence families rather than low intelligence families*.

More broadly still, if it were correct that *high intelligence is a risk factor for autism*, this would offer an insight into why autism is non-adaptive but nevertheless heritable. One answer to the paradox of how common, harmful, mental disorders can nevertheless be heritable is that of *balancing selection*, whereby susceptibility alleles sometimes increase fitness (Keller & Miller, 2006). Assuming high intelligence is adaptive, the autism genotype would then persist in the human population for two reasons: (1) its primary cause is an accumulation of common genetic variants that modulate the severity of synaptic pruning; where polygenic processes convey negative consequences, the risk genes are hard to remove from the population because individually they convey very little risk; (2) high intelligence, which is itself adaptive, is a risk factor for autism. Acting together, *autism would be hard to select out of the human population across evolution because its causes are hard to isolate and its risk factors are themselves adaptive*.

Acknowledgements

This research was supported by ESRC grant RES-062-23-2721 to MT and an ESRC studentship to VK. It is based on work developed under MRC grant G0300188 to MT.

References

Keller, M. C., & Miller, G. (2006). Resolving the paradox of common, harmful, heritable mental disorders: which evolutionary genetic models work best? *Behavioral & Brain Sciences*, *29(4)*, 385-404.

Thomas, M. S. C., Knowland, V. C. P., & Karmiloff-Smith, A. (2011). *A connectionist model of developmental regression in autism (and implications for the wider phenotype).* Manuscript submitted for publication

Thomas, M. S. C., Ronald, A., & Forrester, N. A. (2011). *Modelling associations between levels of description: what can gene-behaviour associations tell us about cognitive process?* Manuscript submitted for publication.