The benefits of computational modelling for the study of developmental disorders: extending the Triesch *et al.* model to ADHD

Fiona M. Richardson and Michael S.C. Thomas

Developmental Neurocognition Laboratory, School of Psychology, Birkbeck College, London, UK

This is a commentary on Triesch *et al.* (2006).

In this commentary, we focus on two aspects of the target article. The first is the decision to separate the gaze following model into two trainable components, a 'When' component to determine when to shift gaze and a 'Where' component to determine where gaze should be shifted to. The second is the lesson that the authors draw regarding multiple causality in developmental disorders, that is, from the finding that very different computational causes can lead to similar deficits in the emergence of gaze following. In that context, we assess a version of the gaze following model given Attention Deficit/Hyperactivity Disorder (ADHD).

In their article Triesch, Teuscher, Deàk and Carlson (2006) implement their *Basic Set* account of gaze following within a computational setting. In doing so, they provide a working parameterized theory with which to further our understanding of the neurocomputational causes of deficits within developmental disorders. A key feature of their model is that the developmental process itself is central to the emergence of atypical behaviour. Through the use of theory-driven parameter manipulations, it is possible to identify atypical precursors that produce later end-state deficits. For example, changing the reward value of looking at faces in different ways in their model is sufficient to characterize gaze following behaviour in both autism and Williams syndrome. This is very much in tune with the neuroconstructivist theoretical approach and empirical studies pioneered by Karmiloff-Smith (1998), in which she proposes that the causes of adult cognitive deficits in genetic developmental disorders must be traced back to their origins in infant precursors for a full understanding of the disorder (see also Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett, 1996).

The biological plausibility of reward-driven learning motivates Triesch *et al.* to select Temporal Difference learning as their central mechanism. This is an embodied approach to modelling in which the agent (in this case, the infant) interacts with a predefined environment, passing from one state to another with each interaction in order to achieve some future reward. The agent's experiences in its environmental setting are balanced through its tendency to either (a) exploit its existing knowledge of states and actions that result in rewards, or (b) explore new states in search of (perhaps greater) reward. In this model, the agent's acquisition of knowledge is controlled through the learning rate, while the decreasing reward for remaining in a given state (modelled here through habituation) encourages the agent to explore new paths of potential reward.

Architectural assumptions

In modelling, simplicity is both a virtue and a necessity. Overly complex models are time consuming to build and run the risk of revealing little about the potential causes of a particular behaviour, since credit and blame assignment can become opaque. The model of Triesch *et al.* is an excellent example of a model that generates a rich set of predictions, despite its simplicity. Some of these stem directly from the model, others from the theoretical framework stimulated by constructing the implementation. Importantly, all aspects of the theory are exposed for scrutiny in a model.

Triesch *et al.* argue that their decision to split the model into two components is not a necessary assumption. One could use a single system but 'learning time would be expected to increase because of the higher dimensionality of the resulting state space' (p. 132). This raises two issues: first, what is the magnitude of the increase in complexity of the learning problem for a single system and second, how does this decision bear on the authors'

Address for correspondence: Fiona Richardson, Developmental Neurocognition Laboratory, School of Psychology, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK; e-mail: f.richardson@psychology.bbk.ac.uk

Table 1 *A state-action table for the When module*

		States (time slices)									
		T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
Actions	Maintain fixation	90	80	55	40	30	18	10	5	3	1
	Shift fixation	10	20	45	60	70	82	90	95	97	99

Note: Values show the system's reward estimates for each action in each state. Values are illustrative and depict the idea that habituation conditions the system to change fixation after some point in time. Note, the action that is selected is determined probabilistically, depending on the system's bias for exploring the environment versus its bias for exploiting its existing knowledge (see Triesch *et al*., equation 7).

Table 2 *A state-action table for the Where module*

		States						
		Where caregiver is looking						
		Region 1	Region 2	Region 3	Region 4	Region 5	Caregiver looking at infant	No info. (not looking at caregiver)
Actions	Fixate 1	80	5	5	5	5	20	10
	Fixate 2	5	80	5	5	5	20	10
	Fixate 3	5	5	80	5	5	20	10
	Fixate 4	5	5	5	80	5	20	10
	Fixate 5	5	5	5	5	80	20	10
	Fixate caregiver	0	0	0	0	0	0	50

Note: Values are illustrative and depict a system that has successfully learned gaze following. Therefore the system gets most reward from looking to the region where the caregiver is looking. Other regions still have reward estimates because the caregiver may be unreliable, the infant may have misperceived direction of gaze, or may chance upon rewarding events by ignoring the caregiver. For states where the infant is already looking at the caregiver, 'Fixate caregiver' reward is zero, since the When table has already requested a change in fixation. The right hand column ('No info.') corresponds to the system's 'sociability', i.e. its tendency to spontaneously look at the caregiver rather than search for an object.

proposal that their model contrasts with innate modularity? Beginning with the complexity issue, a concrete example will help. A reinforcement learning system can be conceptualized as a table that relates states to actions. For each state that the system can be in, the aim is to learn a reward estimate for each possible action.[1] Tables 1 and 2 illustrate possible state-action tables for the When and Where components for a system that has successfully learned to gaze follow in an environment where five regions of space may contain interesting objects and episodes last 10 time steps (note, we invented the values in these tables for illustrative purposes). Separating the system into two components results in a modest total of $20 + 42 = 62$ cell values, for which reward estimates must be learned through social interaction.

Now, consider what would happen if we did not split the model. This would result in a single state-action table, which differs in two ways: (a) we would have to add an extra row to the original Where table to accommodate 'Maintain Fixation', since the current six actions are for cases where fixation changes; and (b) we would need to duplicate this entire table *for every time step*. This would result in a table consisting of $7 \times 7 \times 10 = 490$ cells, for which reward estimates must be learned.

Is learning in this larger system tractable? Does it require more learning events than would be realistically available to the child between, say, 5 months and 10 months, in an 8-hour infant 'working day'? Mitchell (1997) points out that reward learning systems of this sort can be slow to converge, since the same sequence of behaviours must be encountered a number of times to allow reward signals to filter back from the goal state to the earlier steps that must precede it in the sequence. If we follow the authors' relation of model time to real time (1 step = 250 msecs), the learning depicted in their Figure 2 corresponds to around 7 hours of interaction. If a non-modular system could converge in 10 or 100 times this period, this architecture might still be viable, but perhaps not if convergence takes 1000 times as long. Of course, it is undoubtedly premature to relate a model of this simplicity to real-time learning events. Nevertheless, our point is that architectural decisions of this type are non-trivial and, when scalability is taken into account, may form a key assumption of the proposed theory. In this case, we can make the assumption more explicit. The architectural decision corresponds to the

[1] The table can be translated into a two-layer neural network, where the possible states correspond to input units and the possible actions correspond to output units, and the reward estimates correspond to the strength of the connection weight between each input-output unit pair. In this way, Triesch *et al*. relate their model to neural pathways linking the Fusiform Gyrus to Front Eye Fields.

theoretical proposal that the system is granted *a priori* the ability to generalize reward information across time steps.

Interestingly, Triesch *et al.* present their model as an alternative to theories that argue for an *innate gaze following module*. Their use of a modular architecture perhaps belies this status, but more significantly, it is worth noting that one way to define an innate module is to pre-specify proprietary inputs and outputs for a system – even if the system itself uses domain-general processing principles (Thomas & Richardson, 2006). Given that (a) this gaze following system is only exposed to caregiver information, reward information (disconnected from the object that drives it) and time step information, and (b) it is only allowed to drive eye-gaze movements as an output, our view is that Triesch *et al.*'s model comes closer to demonstrating what a realistic innately modular system would look like than opposing innate modularity itself. The model's great bonus, however, is in detailing how a system with these architectural commitments could proactively interact with the environment to acquire its representational content, an element frequently missing from nativist accounts.

Reinforcement learning and developmental disorders: the issue of multiple causality

The second aspect of the model we wish to consider is multiple causality. Triesch *et al.* (2006) state that there has been relatively little use of Temporal Difference learning in modelling development, a point with which we certainly agree. However, in this context, we should mention a recent paper by Williams and Dayan (2005; see also Williams & Dayan, 2004). These researchers use Temporal Difference learning to simulate the developmental profile of impulsivity, based on a model of the role of dopamine in operant conditioning. In this model, the agent (child) must learn to delay an immediate action that gains a small reward in favour of a later action that gains a larger reward. This is a relatively simple model, corresponding only to the When component of the Triesch *et al.* model (e.g. Table 1). Williams and Dayan then go on to demonstrate how manipulating the start-state parameters of their model can capture features of ADHD, specifically an elevated tendency to select actions that achieve immediate rewards rather than those that work towards long-term goals. The Williams and Dayan model leads us to ask, could we generate an 'ADHD' version of the Triesch *et al.* model? What would its gaze following look like? Since Triesch *et al.* invite readers to try out their model, we could not resist the opportunity to find out.

Following Williams and Dayan (2005), there are three ways in which one could simulate ADHD in the Triesch *et al.* model. These are: (1) reduce the learning rate (α in equation 6). This would result in future rewards needing more learning events to percolate back across time steps. Delayed rewards would take more developmental time to exert an influence on immediate behaviour; (2) alter the temperature parameter (τ in equation 7). This would make the model more exploratory, so that it would be more likely to ignore its current knowledge about potential rewards when choosing actions; (3) change the discounting factor (γ in equation 5) so that immediate rewards have a greater influence on learning than future rewards. Figure 1 (a) demonstrates a normal developmental trajectory generated from Triesch *et al.*'s model, along with trajectories for systems with (b) a learning rate reduced by 40%, (c) a temperature increased by 32% and (d) a discounting rate reduced by 38%.

In each case, the model generates the prediction that an 'ADHD' system should show impairments in acquiring gaze following. There is little work on ADHD in infancy as the disorder is not usually diagnosed until childhood. Nevertheless, these simulation results suggest infant precursors of the disorder in gaze following. This finding reinforces Triesch *et al.*'s claim that models of this sort point to multiple underlying causes for behavioural deficits in developmental disorders. But it also raises a methodological issue. Triesch *et al.* use hypothesis-driven parameter changes to the reward signal to produce the developmental trajectories for autism and Williams syndrome. However, the strength of this result must be weighed against the number of possible parameter manipulations that lead to the same outcome. The study of developmental deficits via implemented models requires researchers to explore the parameter space (or background flexibility) of their models (see Thomas & Karmiloff-Smith, 2003; Williams, in press, for methodological arguments). If many parameter manipulations result in the same pattern of atypical behaviour (or a limited set of patterns), it increases the likelihood that the range of deficits is being shaped by the structure of the problem domain rather than the particular parameters that produce sub-optimal learning. This motivates further investigation of the role of the problem domain in constraining development.

In other words, Triesch *et al.*'s two manipulations caused a deficit in acquiring gaze following, but so did three further manipulations derived from a theory of ADHD. One implication is that, in this formulation of the problem, *gaze following* is more vulnerable to developmental disruption than *shifts of fixation to the caregiver*, because gaze following is a sequence involving more steps than caregiver fixation; so it necessarily requires further propagation of the reward signal back to time steps earlier in the sequence.

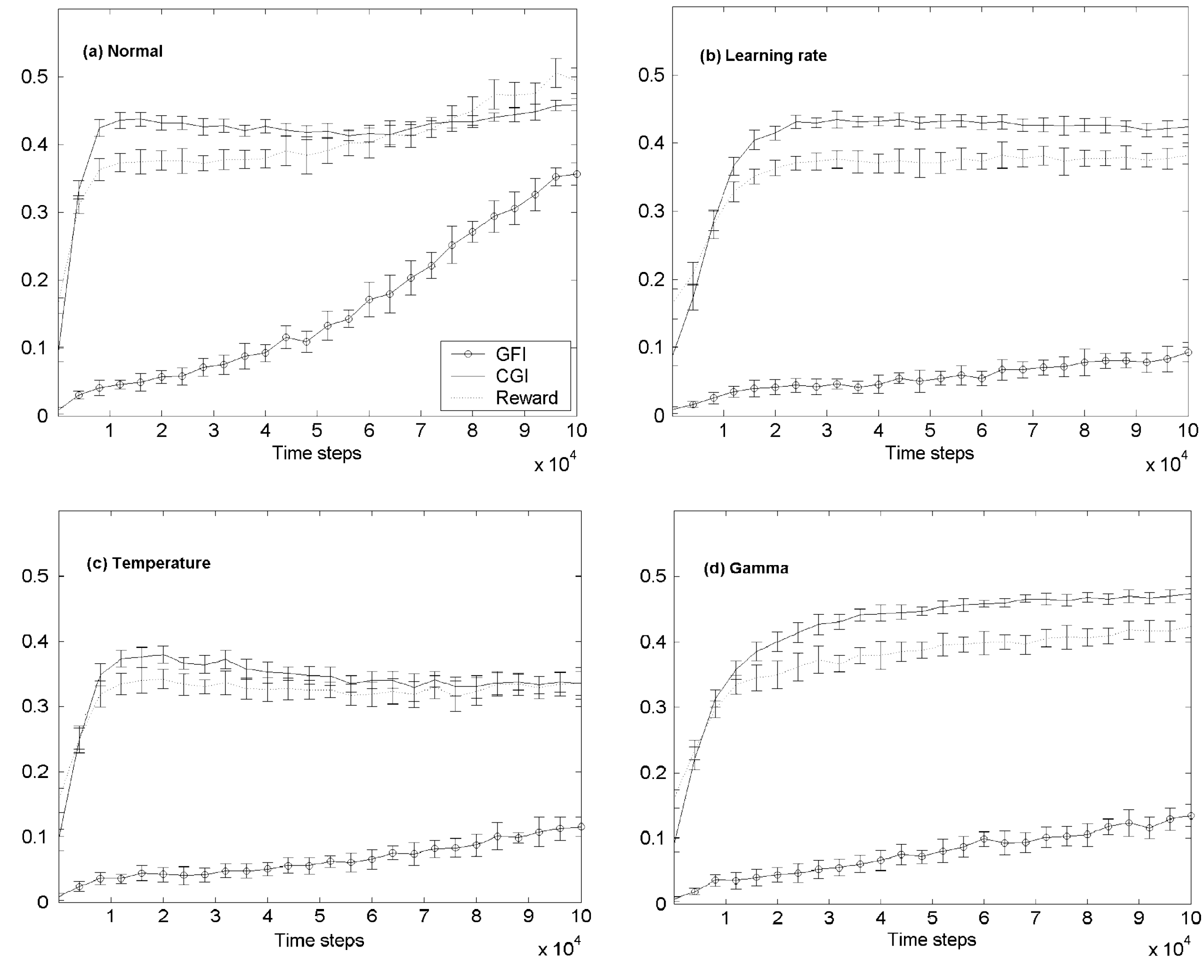

Figure 1 *Emergence of (a) normal gaze following behaviour (the equivalent of Figure 2 in Triesch* et al.*), shown alongside three parameter manipulations, (b) learning rate $\alpha = 0.001$, (c) temperature $\tau = 0.125$, and (d) discounting rate $\gamma = 0.3$, simulating a model with ADHD (default 'normal' parameter values: $\alpha = 0.0025$, $\tau = 0.095$, $\gamma = 0.8$). The caregiver index (CGI), gaze following index (GFI) and reward are shown. The error bars show standard deviations across 15 simulations.*

Importantly, however, if one inspects Figure 1 closely, it becomes apparent that, aside from the much slower and more linear acquisition of gaze following, there are subtle differences between the three atypical trajectories in (b) to (d). For example, the 'exploratory' system in (c) displays declining rates of caregiver fixation and reward, while the 'delayed' system in (b) shows prematurely asymptoting performance on these metrics, and the 'discounted' system in (d) shows persistent gradual improvement. The advantage of implementation in the study of disorders, then, is that we gain a concrete handle on a nebulous notion such as multiple causality, to a point where empirical predictions may be generated. These may allow us to separate sub-types of individuals within a more broadly defined disorder. Using this approach, we recently used a computational model to derive behavioural heuristics, based on test score variability, to distinguish heterogeneous from homogeneous developmental disorder groups, where multiple causality is suspected to be operating (Thomas, 2003).

Summary

In sum, computational models such as that of Triesch *et al.* are not only a powerful tool for understanding the emergence of a particular behaviour, but also force us to face and explore broader issues regarding the structure, organization and mechanisms of learning in cognitive systems.

Acknowledgements

This research was supported by MRC Career Establishment Grant G0300188 to Michael Thomas. We would like to thank the authors of the target article for making their code available to us.

References

Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Science*, **2** (10), 389–398.

Mitchell, T.M. (1997). *Machine learning*. New York: McGraw-Hill.

Thomas, M.S.C. (2003). Multiple causality in developmental disorders: methodological implications from computational modelling. *Developmental Science*, **6** (5), 537–556.

Thomas, M.S.C., & Karmiloff-Smith, A. (2003). Modelling language acquisition in atypical phenotypes. *Psychological Review*, **110** (4), 647–682.

Thomas, M.S.C., & Richardson, F. (2006). Atypical representational change: conditions for the emergence of atypical modularity. In M. Johnson & Y. Munakata (Eds.), *Attention and performance XXI*. Oxford: Oxford University Press.

Triesch, J., Teuscher, C., Deák, G.O., & Carlson, E. (2006). Gaze following: why (not) learn it? *Developmental Science*, **9** (2), 125–147.

Williams, J.O.H. (2005). Problems with models in psychiatry. *Synergies* (Centre for Advanced Studies, Oslo). in press.

Williams, J.O.H., & Dayan, P. (2004). Dopamine appetite and cognitive impairment in Attention Deficit/Hyperactivity Disorder. *Neural Plasticity*, **11** (1), 115–132.

Williams, J.O.H., & Dayan, P. (2005). Dopamine, learning and impulsivity: a biological account of ADHD. *Journal of Child Adolescent Psychopharmacology*, **15** (2), 160–179.