The rise and fall of learning: A neural network model of the genetic assimilation of acquired traits

James R. Watson[‡] and Janet Wiles[◇‡]
[‡]School of Information Technology and Electrical Engineering
[◇]School of Psychology
The University of Queensland, Australia, 4072

{j.watson, j.wiles}@itee.uq.edu.au

Abstract - The genetic assimilation of learned behaviour was introduced to the wider evolutionary computation field by the classic simulation of Hinton and Nowlan. Subsequent studies have analysed and extended their initial framework, contributing to the understanding of the often counterintuitive relationship between evolution and learning. We add to this increasing body of literature by presenting an evolving population of neural networks that plainly exhibit the Baldwin effect. Phenotypic plasticity, embodied in the literal learning rate of the neural networks, is evolved along with the network connection weights. Significantly, high levels of plasticity do not cause the population to genetically stagnate once correct behaviour can be learned. Rather, continuing inter-population competition drives the levels of learning down as beneficial behaviour becomes genetically specified. By observing the evolving learning rate of the agent population, and by comparing learned and innate agent responses, we demonstrate the Baldwin effect in its entirety.

I. Introduction

The relationship between inter-generational processes like evolution and phenotypic traits such as learning has been of interest to fields as diverse as biology, psychology and evolutionary computation. Since such a relationship necessarily emerges over many generations, computational models of evolution and learning provide an effective and convenient way to study these phenomena.

The Baldwin effect, first described separately by Baldwin [2], Morgan [17] and Osborn [18], accounts for the apparent Lamarckian tendency of learned behaviours to become genetically specified – within the constraints of accepted Darwinian evolution. Initially confined to the field of biology (e.g., Waddington [21]), the effect was introduced to the wider evolutionary computation community by the simulations of Hinton and Nowlan [10].

This paper inspired much interest in the relationship between learning and evolution (at least within the evolutionary computation community – genetic assimilation has not always been accepted in biology due to suspicions that such phenomena were Lamarckian [19]). Subsequent works have involved the addition of cultural learning [3], the evolution of neural network learning rates [4] and the evolution of associative learning [20] to name a few. However, understanding the original focus of Hinton and Nowlan's simulation – the Baldwin effect – is still a topic of discussion. It has been studied with a classic population genetics approach [1, 6] and with boolean neural networks based on grammar tree encodings [8]. Studies concerning the influence of evolutionary operators on genetic assimilation [11, 12] as well as comparisons between Lamarckian and Baldwinian forms of evolution [22] have also been undertaken.

But the fundamental issue of why genetic assimilation of learned behaviour occurs at all *within the constraints of Darwinian selection* has come down to a balance between the benefit of being able to learn, and the corresponding costs involved with such behaviour [13, 15]. In Hinton and Nowlan's simulation, plasticity was required to find the solution in the first case. But this flexibility was balanced against a fitness function that preferred solutions in proportion to how genetically correct they were. This cost of learning was sufficient for the population to move, over time, towards a more genetically correct solution. Thus, the genetic assimilation of *acquired* behaviour was demonstrated. The entire process is illustrated in Figure 1. Our replication of Hinton and Nowlan's results show just how quickly the population is able to learn the solution, as is shown by the addition of a 'correctly learned alleles' line.

This paper illustrates the same idea occurring under a different guise. We simulated a population of evolving neural networks, where both the network weights and the individual learning rates were allowed to evolve. In keeping with the Hinton and Nowlan analogy, correct behaviour was initially difficult to achieve in the absence of learning (thus initial populations with high phenotypic

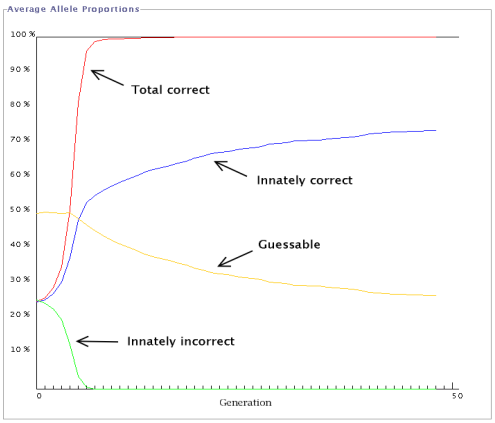

Fig. 1. Replication of Hinton and Nowlan's original simulation, with the addition of a 'learned correct' line. The population is able to quickly learn the solution as innately incorrect alleles are removed from the population. But the number of guess-able alleles remains relatively constant until practically all of the population is able to *learn* the solution – which is when selection focusses on more innately correct solutions.

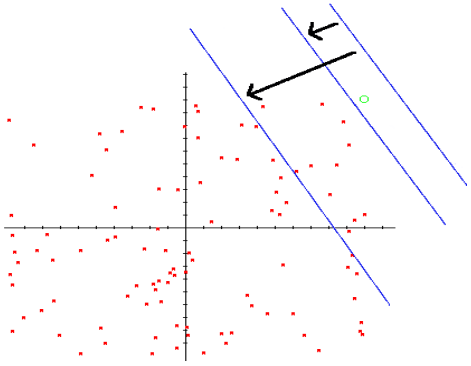

Fig. 2. Two-dimensional representation of simulation hyperspace. Food is represented by the circle in the upper-right corner, with toxin (dots) occupying the rest of the space. Black arrows illustrate the effect a large (large arrow) and a small (small arrow) learning rate can have while an agent is trying to differentiate food from toxin. This amplification of phenotypic noise is the primary cost of the form of plasticity used in our simulations.

plasticity were expected). Since the learning rate effectively amplifies an agent's response, the cost of learning was the introduction of noise in phenotypic behaviour. For example, network output could be continuously over-corrected if the learning rate is too large (see Figure 2).

French and Messinger [7] have also investigated the Baldwin effect using populations of agents that could evolve their own phenotypic plasticity. Within their framework, they found that once correct behaviour could be learned, genetic movement towards correctness stagnated. This lack of genetic assimilation was in the absence of a cost to learning. Our aim in this paper is to show the complete genetic assimilation of learned behaviour (i.e., the rise and fall of population plasticity) using a population of neural networks evolving their own learning rate.

II. The Baldwin Effect

As previously mentioned, the Baldwin effect describes a potential way learned behaviours can become genetically encoded over the course of evolution, within the (biologically realistic) constraints of Darwin's theory of evolution [5]. The significance of the Baldwin effect lies in the assertion that individuals (phenotypes) cannot directly influence their genetic composition (genotypes). Operating in environments in which individual plasticity has both benefits and costs, the Baldwin effect (applied to learning) can be best conceptualised as comprising two consecutive stages (illustrated in Figure 1):

(i) the ability to learn a task gives an individual (or individuals) selective advantage over the rest of the population, causing subsequent generations to become increasingly dominated by that individual(s)'s genes, and

(ii) the costs of learning cause selective pressure to favour individuals who, because of recombination and/or mutation, are genetically predisposed such that they don't have to learn as much as their peers. Over time (often many generations), the ability becomes more and more genetically specified.

However, empirical investigations into this process have found that the "genetic assimilation of acquired traits" [13] is not as simple a process as may be initially thought. Hinton and Nowlan's original framework [10], rarely, if ever evolves the complete genetic specification of the solution [9]. Other factors such as the cost of learning [13], the amount of phenotypic plasticity [7], mutation rate [6] and selection algorithm [23] (for example) have been shown to be crucial factors governing the relationship between learning and evolution, and subsequently the existence of the Baldwin effect.

In other words, the Baldwin effect can be seen as a phenomenon sensitive to multiple parameters, requiring the modeller's careful consideration when choosing all aspects of a simulation. This susceptibility to simulation design is due to the tendency of modelling decisions (such as selection algorithm, mutation rate, fitness function, even population size), to affect the nature of the task and hence the fine balance between the benefits and costs to

individual learning. As well noted by Mayley [13], too much benefit and not enough costs involved with learning, and there will be insufficient selective difference between phenotypes that need to learn a little and those that need to learn a lot. Conversely, too little benefit and/or too great a cost, and the task will never be learned. In effect, the simulation parameters need to be tuned so this balance is emphasised. This notion was an inherent characteristic of Hinton and Nowlan's simulation, and balancing these costs defined the initial point of our simulation.

III. Simulation Framework

The simulation consisted of a population of single-layer feed-forward neural networks. Each network had n input units and a single output unit (whose output value was determined by a sigmoidal activation function). The connection weights between the input units and the output unit, and the value of the learning rate, formed each agent's genotype.

The environment consisted of food and toxin input strings comprised of a combination of n -1's and 1's. If an agent's response to an input string was higher than a fixed threshold, that agent was deemed to have 'eaten' it. A food counter was incremented if food was eaten, and decremented if toxin was eaten. The fitness function was based on the relative amounts of food versus toxin consumed. The food representation was fixed (i.e., a constantly-defined correct response was specified) while the toxin representation was constantly changing. (It was due to the changing toxin representation that the input strings did not consist of 0's and 1's. This combination would have introduced unnecessary bias into the simulation by hiding potentially large network weights connected to 0's until the representation was altered. Such a situation could lead to agents exhibiting dramatic changes in fitness, with the changes being an unwanted artifact of the choice of representation). Food and toxin representations always remained mutually exclusive.

Each agent lived for a fixed number of 'days', each day consisting of

(i) evaluation of agent output given the current input / update the food counter
(ii) using the delta learning algorithm, agent weights were updated to bring the next output closer to what was expected for the current representation (i.e. below or above the threshold), and
(iii) the current toxin representation was randomly changed

Tournament selection (as described in Mitchell [16]), was used to create each successive generation, treating agents with the higher food counts at the end of the fixed number of agent days as being of higher fitness.

Neural network populations were asexual, and selected agents reproduced by passing on their inherited connection weights (*not* the learned weights) to the next generation. The network connections, the bias and the agent learning rate were mutated with a small chance during this process.

A. Poised on the Brink

Several steps were taken to make the simulation conform to the conditions identified in Section II – that is, the task was made just hard enough so that a couple of the initial agents should learn it, and the cost of learning such that sufficient selection pressure to favour agents that required less learning was present.

In order to create the conditions required to satisfy the original stage of the Baldwin effect – a population's transition from being unable to perform the task to being able to accomplish it with learning – a series of simulation runs with varying parameters were performed to determine the combination of parameters that poised the population on the brink of discovery. The population size was fixed, and the noise of the simulation (mutation rate, rate of toxin changes) predetermined to give a combination of parameters that served as a base for the ensuing trials. Agent responses over time were measured using this framework. A specific output threshold and lifetime were subsequently chosen from these trials so that, like Hinton and Nowlan, correct behaviour was just within reach of one or two of the best members of the initial agent population.

B. Measuring Genetic Change

It is possible for evolutionary change to still be occurring even after phenotypic behaviour has (apparently) stagnated. In fact, this is precisely what happens during the 'second stage' of the Baldwin effect – a population that behaves correctly with learning evolves over time into a population that behaves correctly with less of it (see Figure 1). As a consequence, an analysis of phenotypic behaviour is inadequate in determining how and when genetic assimilation occurs.

Therefore, measurements of genetic performance are required in order to observe the relationship between learning and evolution. The genetic encoding used in our simulation, coupled with phenotypic learning, meant that there was no way of directly correlating specific genes and phenotypic traits. (That is, emergent behaviour is the result of interactions between innate network weights and

learning). In addition, although fixation of the food representation meant that there was one correct response when an individual was presented with food, many possible genotypes existed that could satisfy this phenotypic requirement. Consequently, measurements of genetic performance could not rely on the literal values of the phenotype or genotype networks.

Two indirect methods of measuring genetic change were employed to detect the existence of the Baldwin effect. The first was to measure the output of both the innate (genotypic) and learned (phenotypic) networks and observe the differences between the two responses. Since learning was the only process capable of modifying an agent's behaviour during a lifetime, any difference between these two measurements indicate the application of phenotypic plasticity. The second method was to observe the values of the evolved learning rate, and to compare them with phenotypic performance. If phenotypic performance did not follow a course parallel to the evolving learning rate, then genetic changes could be inferred.

IV. Results

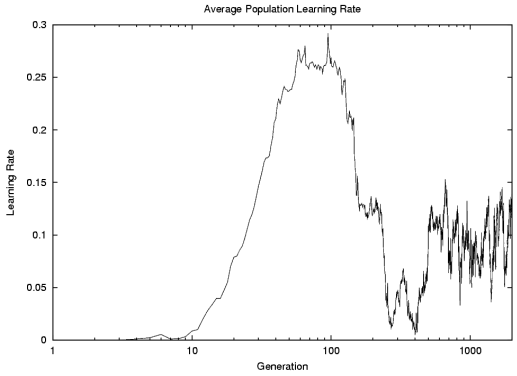

Fig. 3. Evolved learning rate. The Baldwin effect can be seen in the initial rise then fall of the learning rate. The initial steep rise is correlated with improvements in the (phenotypic) consumption of food. The subsequent decline, *which did not correspond with any deterioration in phenotypic behaviour* (see Figure 4), indicates populations evolving to require less learning. Note that the learning rate never completely reduced to zero, consistent with the finding of Hinton and Nowlan [10], Harvey [9] and others.

The population's average evolved learning rate is shown in Figure 3. The two classic stages of the Baldwin effect can be clearly seen by comparing Figure 3 with Figure 4. Each graph represents the respective population averages. The learning rate rises as selection favours those that can learn the solution. At approximately the hundredth generation (when the whole population has practically learned the task), the average learning rate sharply declines as selection favours – from among individuals that can all learn the solution – agents that incur less learning costs in the course of a lifetime.

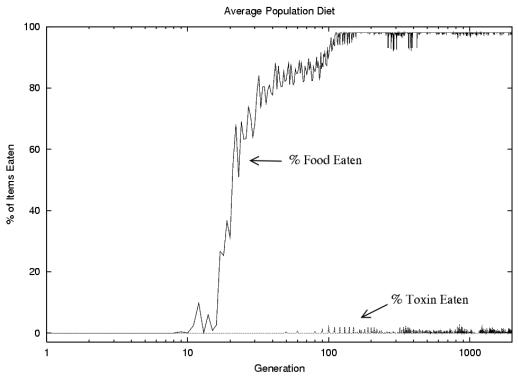

Fig. 4. The average percent of food and toxin eaten, which shows how quickly the agents were able to learn perfect behaviour. Compared with Figures 5 and 3, it can be deduced that genetic movement towards more innately correct solutions evolve once the cost of learning drives competition between agents that have learned the task.

The population's transformation from being a learnable (but costly) distance from the correct phenotype to being genetically correct is never complete (as shown by the learning rate hovering between 0.025 and 0.15 after the initial rise and fall). One explanation for this remaining plasticity is that selection is given increasingly diminishing grounds to distinguish between agents that can learn the solution, as the average cost of learning diminishes the closer the genotypes approach correct behaviour. Effectively, the conditions which poised the simulation to exhibit the Baldwin effect are lost before complete genetic assimilation can occur. Hinton and Nowlan's simulation also exhibited such residual learning, which they attributed to a lack of selection pressure once there was only a small amount of learning required. Harvey [9] explains this phenomenon as being caused by hitchhiking and genetic drift. But such an effect is not inherent to the Baldwin effect. Residual learning has been found to be heavily dependent on the choice of selection algorithm [23]; accompanying fitness proportionate selection (the algorithm used by Hinton and Nowlan), but practically non-existent with Elitist Tournament selection (used in these simulations). Thus, the residual learning characterised in our simulation is also likely to be due to environmental noise (mutation and changing toxin representation), as it has been shown that asexual populations tend to be less resistant to these environmental pressures than sexual ones [11]. The emergent

behaviour of a sexual population in our simulation remains an open question.

A. Learning versus Evolution

As a final measure of genetic assimilation, we consider the difference between the learned and innate food response. The innate food response is the network output that results from an agent's *inherited* weight values (and is used solely for comparison purposes). As mentioned, the learning rate was the only means by which an agent's response could be modified during its lifetime. Therefore, any differences between the actual response value given and the innate response value is entirely attributable to phenotypic plasticity. The difference between the learned and innate response in Hinton and Nowlan's simulation is shown in Figure 1 as the difference between the 'total correct' and 'innately correct' lines. A similar comparison for our neural network learners is shown in Figure 5.

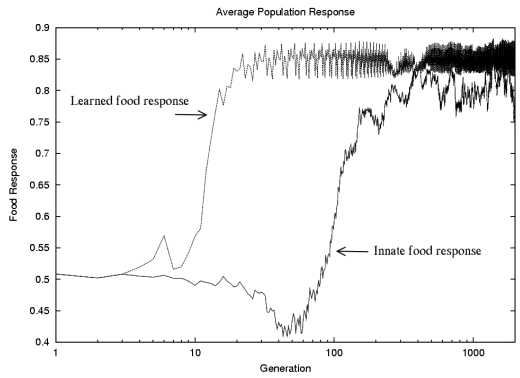

Fig. 5. The average innate and learnt food responses of agents over time (note the logarithmic scale). The first stage of the Baldwin effect can be seen in the rapid increase in the population's ability to learn to output a threshold that chooses food (0.8 in this case). The second stage is illustrated by the rapid rise in the innate response, which meant less and less learning was involved.

Figure 5 clearly shows the average population quickly acquiring the solution through plasticity early in the simulation (learnt response), as individuals that were incapable of learning beneficial traits were selected out. Genetic assimilation of this acquired behaviour occurs soon after (innate response). Once again, this evidence suggests that it is not until the entire population is comprised of good learners that the cost of learning outweighs its benefits, and selective pressures favour genetic assimilation.

Both learnt and innate responses continued to fluctuate after the two stages of the Baldwin effect had been observed, since the frequently changing toxin representation and mutation continued to apply environmental noise. Agent response to toxin is not shown. Both innate and learned toxin responses exhibited very noisy behaviour, due to this environmental noise. The learned toxin response followed the general path of the innate response, but was subject to greater variance in relation to the size of the learning rate (shown in Figure 3). Toxin responses very rarely crossed the choice threshold.

Interestingly, as the evolved learning rate climbed to compensate for the population's distance from the correct phenotype, a corresponding overall *decrease* in genetic 'correctness' was observed. In addition to the fact that *average* innate response dropped to levels below what was expected if the genetic composition was random (a sigmoidal output of 0.5), this observation occurred in too many simulation runs to be written off as genetic drift (compensated for by a high learning rate). It is more likely to be a consequence of the learning rate itself. In periods of high plasticity, the neural network agents would benefit from having small weights that allow a large learning rate finer-grained solution search. Large weights at this stage would simply hamper a big learning rate, causing learned responses to bounce between one incorrect side of the solution hyperspace to another.

B. Population Size

As an interesting aside, it was found (while varying population size) that small populations with learning (even mutant-champ systems) were capable of phenotypically finding the solution *as quickly* as very large populations (of up to 200,000 individuals). Given that the time to run the simulations is approximately linearly proportional to the size of the population, optimization tasks that can disregard phenotypic versus genotypic considerations could benefit from focusing on agent plasticity rather than trying large population sizes to search the fitness landscape.

V. Conclusions

Two methods of measuring genetic change over time were introduced in this paper. The first technique was to compare actual (i.e., after learning) network output with innate (i.e., inherited) responses. The second was to compare the values of the evolved learning rate with actual phenotypic performance. These methods used to observe the Baldwin effect were necessarily distinct from directly measuring actual genetic values (as was done by Hinton and Nowlan, but meaningless in the context of

our simulations), or agent fitness in isolation (which is known to be an inaccurate measure of genotypic trends [15]).

Using these methods, we have successfully demonstrated the Baldwin effect in a population of neural network agents evolving individual learning rates. In addition to evolving increasing levels of plasticity to acquire correct behaviour, the explicit cost of learning in our fitness function caused agent plasticity to drop and the genetic assimilation of this learned behaviour to occur. This balance between the costs and benefits of learning was found to be the fundamental factor influencing the existence of the Baldwin effect (supporting Mayley [13, 14]). Factors that have been previously identified as impacting the genetic assimilation of acquired behaviour are likely to all do so by altering this balance.

Residual learning, which has been noted in other studies of the Baldwin effect [9, 23] was found to be due to the environmental noise of the simulation. In addition, a drop in genotypic performance was observed when plasticity was high (and the cost of this flexibility was low). This drop was because a diminished genetic influence allowed learning finer-grained search of the problem space.

The results presented in this paper add to the increasing body of work acknowledging forces acting at the genetic level, but existing at the much higher level of phenotypic behaviour. Far from being Lamarckian in nature, the concepts of genetic assimilation draw to attention the complex interplay between many seemingly disparate processes occurring over evolutionary time-scales. Although such artificial simulations can not claim a direct analogy with the biological world, they are providing a clearer understanding of the *possible* interactions between learning and evolution.

VI. Acknowledgements

This study was supported by the University of Queensland Graduate School (JRW) and the Australian Research Council (JW).

References

[1] Anderson, R. W. (1995). Learning and evolution: A quantitative genetics approach. *Journal of Theoretical Biology*, 175:89–101.

[2] Baldwin, J. M. (1896). A new factor in evolution. *American Naturalist*, 30:441–451 536–553. Reproduced in Belew, R. K. & Mitchell, M. (Eds.), *Adaptive Individuals in Evolving Populations*. Addison-Wesley, Reading, MA.

[3] Belew, R. K. (1990). Evolution, learning and culture: computational metaphors for adaptive search. *Complex Systems*, 4(1):11–49.

[4] Chalmers, D. (1990). The evolution of learning: An experiment in genetic connectionism. *Proceedings of the 1990 Connectionist Summer School*, pages 81–90.

[5] Darwin, C. (1859). *The Origin of Species by Means of Natural Selection*. The Modern Library.

[6] Fontanari, J. F. and Meir, R. (1990). The effect of learning on the evolution of asexual populations. *Complex Systems*, 4:401–414.

[7] French, R. and Messinger, A. (1994). Genes, phenes and the Baldwin effect: Learning and evolution in a simulated population. In Brooks, R. A. and Maes, P., editors, *Artificial Life IV*, pages 277–282.

[8] Gruau, F. and Whitley, L. D. (1993). Adding learning to the cellular development of neural networks: Evolution and the Baldwin effect. *Evolutionary Computation*, 1(3):213–233.

[9] Harvey, I. (1993). The puzzle of the persistent question marks: A case study of genetic drift. In Forrest, S., editor, *Proceedings of the Fifth International Conference on Genetic Algorithms*, pages 15–22, San Mateo, CA. Morgan Kaufmann.

[10] Hinton, G. and Nowlan, S. (1987). How learning can guide evolution. *Complex Systems*, 1:495–502.

[11] Jones, M. and Konstam, A. (1999). The use of genetic algorithms and neural networks to investigate the Baldwin effect. In Carroll, J., Hiddad, H., Oppenheim, D., Bryant, B., and Lamont, G. B., editors, *Proceedings of the 1999 ACM Symposium on Applied Computing*, pages 275–279.

[12] Ku, K. W. C. and Mak, M. W. (1998). Empirical analysis of the factors that affect the Baldwin effect. In Eiben, A. E., Bäck, T., Schoenauer, M., and Schwefel, H.-P., editors, *Parallel Problem Solving from Nature – PPSN V*, pages 481–490, Berlin. Springer.

[13] Mayley, G. (1996a). The evolutionary cost of learning. In Maes, P., Mataric, M. J., Meyer, J.-A., Pollack, J., and Wilson, S. W., editors, *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive behavior*, pages 458–467. MIT Press.

[14] Mayley, G. (1996b). Landscapes, learning costs and genetic assimilation. In Turney, P., Whitley, D., and Anderson, R., editors, *Evolution, Learning and Instinct: 100 Years of the Baldwin Effect*, Cambridge, Ma. MIT Press.

[15] Mayley, G. (1997). Guiding or hiding: Explorations into the effects of learning on the rate of evolution. In Husbands, P. and Harvey, I., editors, *The Proceedings of the Fourth European Conference on Artificial Life*.

[16] Mitchell, M. (1996). *An Introduction to Genetic Algorithms*. MIT Press, Cambridge, MA.

[17] Morgan, L. C. (1896). On modification and variation. *Science*, 4:733–740.

[18] Osborn, H. F. (1896). Ontogenic and phylogenic variation. *Science*, 4:786–789.

[19] Smith, J. M. (1987). When learning guides evolution. *Nature*, 329:761–762.

[20] Todd, P. M. and Miller, G. F. (1991). Exploring adaptive agency II: Simulating the evolution of associative learning. In Meyer, J. and Wilson, S., editors, *From animals to animats: Proceedings of the First International Conference on Simulation of Adaptive Behaviour*, pages 306–315. MIT Press, Cambridge.

[21] Waddington, C. H. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, 150:563–565.

[22] Whitley, D., Gordon, V. S., and Mathias, K. (1994). Lamarckian evolution, the Baldwin effect and function optimization. *Lecture Notes in Computer Science*, 866:6–15.

[23] Wiles, J., Schulz, R., Bolland, S., Tonkes, B., and Hallinan, J. (2001). Selection procedures for module discovery: Exploring evolutionary algorithms for cognitive science. In Moore, J. D. and Stenning, K., editors, *Proceedings of the 23rd Annual Conference of the Cognitive Science Society (CogSci 2001)*, pages 1124–1129. Lawrence Erlbaum Associates.