Selective attention in linked, minimally cognitive agents

Robert Ward (1)

Ronnie Ward (2)

(1) Wolfson Centre for Clinical and Cognitive Neuroscience; University of Wales, Bangor, UK

(2) Dept. of Computer Science; Texas A&M University

Correspondence to:
Robert Ward
Wolfson Centre for Clinical and Cognitive Neuroscience
School of Psychology
University of Wales, Bangor
United Kingdom LL57 2AS
r.ward@bangor.ac.uk

Abstract

We investigate minimally cognitive agents for attention and control of action. VA (visual agent) can attend to one target while ignoring another, and then reallocate processing. Here we review some of our previous results with VA, showing how an analysis of a minimally cognitive agent can both support previous ideas of selective attention and action, as well as suggest new possibilities. We investigate a simple case in which multiple agents are combined to solve tasks that no single agent can complete on its own.

Cognitive agents and their environment interact within a continuous perception-action cycle: perceptions guide actions, actions alter the environment, and the altered environment in turn generates new perceptions (Gibson, 1979). However, computational models in psychology have tended *not* to focus on the dynamic nature of the perception-action loop. Even models of perception-action linkage are usually static in nature: processing begins with the presentation of static stimuli, and ends with the selection of an action which has no effect on the stimulus environment (Cohen, Dunbar, & McClelland, 1990; Schneider, 1995; Ward, 1999). In many cases, action is not modelled seriously, perhaps consisting of an activated output node representing one of several discrete alternative responses. In this respect, the models largely follow the available empirical work. Although much is known about the capacity, time-course, and brain regions important for selective attention and action, in the large majority of studies, the responses measured have little to do with the dynamic nature of the perception-action cycle. Usually, participants make some discrete response, like a keypress, and the trial ends with no impact on the agent's environment.

Theories and models generated for static situations and discrete responses may be inadequate for more interesting tasks and realistic agent-environment interaction. We want to push psychological models towards more complex forms of action, and to encourage thinking about how selective attention and action work within the perception-action cycle. In this respect, a useful class of models, sometimes called embodied, situated, and dynamical (ESD) agents, stress what Clark (1999) calls "the unexpected intimacy between the brain, body, and world". The ESD approach emphasizes the contextually-bound nature of solutions to cognitive problems, and therefore considers specific tasks in their entirety, rather than isolated but general-purpose attention, perception, or motor subsystems. In ESD agents, control of action does not arise through general-purpose representation, but emerges from a distributed system involving the brain, body, and world.

An important form of the ESD approach for our purposes is the “minimally cognitive” agent, developed by Beer (1996). “Minimally cognitive” agents are intended to be tractable models that are embodied, situated, dynamic, and that can do interesting cognitive tasks. For example, Figure 1 shows the visual agent (VA) developed by Slocum et al (2000), evolved by genetic algorithm to selectively attend and respond to targets in its 2D environment. The agent was equipped with an array of proximity sensors, left and right motor units to control its movement, and a small hidden layer of eight units. The targets “fell” from the top of the 2D environment, and VA was meant to “catch” them by moving under them before impact. In our simulations, the two targets fell straight down, and with constant velocity, the first target (T1) by definition having greater velocity than the second target (T2). Following Slocum et al (2000), the agent was controlled by a continuous-time recurrent neural network (CTRNN), with a reasonably generic network architecture (see Figure 1)..

Although catching one target is trivial, catching two targets is definitely not. Catching two targets requires numerous cognitive operations for success, including: (1) correctly prioritizing one of the two targets for immediate action; (2) before impact, tieing motor systems to the movements of the selected target, and insulating them from the effects of the other target; and (3) after catching the first target, the second target, which was previously insulated from the motor system, must now be allowed to control the motor system. These processes of selection, response control, and reconfiguration following a change of targets are all important themes in current selective attention research (for an overview, see Driver, 2001). Not only does the task require interesting cognitive abilities, but it is also a good example of an embodied agent, embedded within a genuine perception-action loop: action involves an extended sequence of motor commands in response to sensory inputs: the senses guide action, and action generates new sensory inputs.

VA therefore seems like a useful starting place for looking at a new class of models, that extend beyond the limits of static situations, free of any meaningful agent-environment interaction. We have then recently undertaken *psychological* analyses of this agent. Our aim was to understand the bases for selective attention and control of action within the agent. Our approach might be fairly described as “artificial comparative psychology”. We are seeking to understand the psychological principles that allow VA to control its attention, and then compare to what is found in the human attention literature.

We reasoned that, as long as VA is doing a *genuine* selective attention task, two outcomes are possible, and both are interesting. First, it might be that VA operates according to new principles not yet seen in the human literature. In this case, VA acts as a model to new experimentation and theory. Second, it might be that the agent operates according to models already developed in the human literature. In this case, VA serves as a tractable, embodied model system. However, the validity of either outcome depends upon the validity of the selective attention, an issue we address in greater detail elsewhere (Ward & Ward, in press).

The psychology of attention in a minimally cognitive agent

In fact, our analyses of VA uncovered both established and novel concepts in the human attention literature. Here we briefly review our analyses of VA, describing its psychological mechanisms for attention and control; however, complete details are available elsewhere (Ward & Ward, in press; Ward & Ward, 2006).

Reactive inhibition

It can seem effortless to switch attention from one object to the next, but this is probably an illusion. Robust costs are found whenever people have to switch their attention, particularly when reallocating attention to an object that was previously being ignored. These costs are large enough and reliable enough that the phenomenon as been given its own label. *Negative priming* refers to costs observed when a previously ignored object must be subsequently attended and reported. Models of *reactive inhibition* (Houghton, & Tipper, 1994) propose that negative priming is a direct reflection of the competitive strength of the object for attention. According to reactive inhibition accounts, a target is selected by inhibiting the nontarget items. The more salient the nontarget -- that is, the more powerfully it tends to attract attention -- the greater the inhibition needed to prevent the object from gaining control of behavior. The flipside is that if attention needs to be subsequently allocated to this inhibited item, it will take time to overcome this inhibition. Thus, more negative priming is observed for highly salient nontargets than for less salient ones (Grison & Strayer, 2001). The otherwise unintuitive result -- that the *more* salient an ignored item, the *harder* it is to switch attention to it -- follows in a straight-forward way from reactive inhibition.

We found that VA attended to the first target (T1) through reactive inhibition of the second (T2), as demonstrated through a variety of analyses. One example is the time required to reallocate processing from T1 to T2. After T1 impacted, it was cleared from the environment, so that it could no longer activate VA's sensor array. After catching T1, we noticed a clear hesitation before VA moved again to catch T2. We measured this hesitation as a function of the salience of T2. On some trials, the speed and spatial separation of the falling targets meant that a good catch of T1 would leave T2 outside the sensors. We called these out-of-view (OOV) trials. On OOV trials, after catching T1, VA would be looking at a simply blank sensor array. For efficient performance on OOV trials, VA would need memory for the position of T2. In fact, memory for T2 position was well above chance. VA moved directly towards the location of the unseen target on about 85% of novel test trials.

However, what was interesting from the perspective of attention and control of action was that hesitation varied according to the visibility of the T2 item, as shown in Figure 2. On a sample of 500 novel trials, hesitation on OOV trials was reduced compared to trials in which T2 was in view, $F(1,498)=95.9$, $p<0.0005$. This finding follows cleanly from the idea of reactive inhibition. The more salient in-view T2's required more inhibition than the OOV ones. Therefore, when it was time to reallocate attention to T2, there was more inhibition to overcome.

Overall, we found four lines of evidence for reactive inhibition in VA. First, as already mentioned VA was slower to respond to more salient, compared to less salient T2s. Second, we isolated a signal within VA's hidden layer that inhibited responses towards salient T2s. Third, we found that inhibition of T2 depended upon the degree of conflicting motor activations within VA. Fourth, we observed the opponent-process loops within VA's evolved network. These results all pointed to an internal, inhibitory signal based on T2 salience (details available in Ward & Ward, in press).

We also found some similar results in experiments with people. We presented observers with a simulation of the two falling targets, using stimulus parameters essentially identical to what we used with VA. Before the trial began, the two balls appeared on the screen. After the trial was initiated, the balls began to “approach” the observer, increasing in size. Participants could move the viewing window left and right by pressing the arrow keys, like a first-person shooter game, and they were instructed to “catch” the targets as accurately as possible by centering the viewing window on each target at the time of impact.

After catching T1, there was a clear period of hesitation, approximately 120 ms, before the next movement towards T2. Human participants, just like VA, were slower to respond towards in-view compared to OOV T2s. We also found that people were also slow to respond to T2's made salient by other means, including color contrast (Ward & Ward, in press). In this case, one target was nearly the same color as the background, and the other was a very salient bright red. T2 could either be relatively salient (T1=background color; T2=red), or unsalient (T1=red; T2=background color). In this experiment, the stimulus parameters were set so that there were no OOV trials. The simple question was how would hesitation vary as a function of T2 salience. As predicted, hesitation to respond to T2 increased when T2 was relatively salient. Again, this otherwise counter-intuitive result follows straightforwardly from a reactive inhibition account.

Distributed, not localized, conflict monitoring

One way to think about the effect of the two targets on VA's behavior is in terms of *cognitive conflict*. Cognitive conflict occurs when neural pathways associated with different concurrent processes interfere with each another (Botvinick, Braver, Barch, Carter, & Cohen, 2001). In psychology, the Stroop task is perhaps the best studied example of cognitive conflict. In this case, the identity of a colored word can interfere with color naming (Mari-Beffa, Estevez, & Danziger, 2000). The two-target catching task is another example of a conflict task. VA is most clearly in a state of cognitive conflict when it is located between the two targets. In these circumstances, T1 will be signalling a move in one direction, and at the same time, T2 will be signalling a move in the opposite direction.

So how is cognitive conflict managed? Botvinick et al (2001) proposed a two-part “evaluate-regulate” approach. The top-down conflict-monitoring system first detects conflict in underlying neural structures. Control mechanisms are then invoked which regulate processing in some appropriate way to reduce conflict. Botvinick et al suggest that conflict monitoring is localized to a dedicated module, argued to be the anterior cingulate cortex

(ACC) in humans. In support of this claim, Botvinick et al (2001; Botvinick, Cohen, & Carter, 2004) review a variety of evidence showing ACC activation during conditions producing response conflict.

We therefore examined VA for evidence of a localized conflict monitoring mechanism. For VA, we did not observe any set of units dedicated to monitoring or regulation. Instead we found a system that used competitive balancing of response biases, and distributed conflict management (Ward & Ward, 2006). This does not mean that other types of systems might use a dedicated module for conflict detection, and in fact, in the next section, we use just such a dedicated module. However, the current simulation does demonstrate that a model system operating within a genuine perception-action cycle, and facing real problems of response conflict (and solving them!) does not require this kind of dedicated module. It is worth remembering that there was no module for either conflict detection or regulation built-in to the VA architecture. Instead the computational properties of the catching task shaped the solution.

Further investigation is still needed to understand how conflict is managed by VA. Yet these results might still help us understand current data regarding the ACC. It is well-established that ACC is activated during conflict tasks, but the significance of this activation is currently being dabated. For example, neurological patients with damage to ACC can be unimpaired in response conflict tasks (Fellows & Farah, 2005). ACC activation might therefore be a correlate rather than causal factor of performance in response conflict tasks, such as arousal (Critchley et al., 2003), or effort (Mulert, Menzinger, Leicht, Pogarell, & Hegerl, 2005). Our results with VA, and its success in resolving cognitive conflict without explicit conflict monitoring, lead us to likewise question the causal link between ACC activation and the resolution of response conflict (Ward & Ward, 2006). The ACC may be one of many sites in which there is competition such as we observed in VA, biased activation for different actions.

Emergent abilities in combined agents

Let's assume for exposition's sake that the human brain has on the order of 100 billion neurons (about 20 billion cortical neurons (Pakkenberg & Gundersen, 1997), and about 100 billion cerebellar neurons (Andersen, Korbo, & Pakkenberg, 1992)), each one of which has far more sophisticated processing capabilities (Graham & van Ooyen, 2004) than the 17 simple CTRNN units we used with VA. Further, there are an estimated 240 *trillion* synapses in the cortex alone (Koch, 1999), compared to 170 synapses for VA. This is a big gap. So, how do we scale up from a minimally cognitive agent to a brain? Which direction do we go? With more VAs, or with bigger VAs? In other words, is the human brain's cognitive architecture best thought of as a relatively small number of large, multi- or general-purpose networks, or as a collection of billions of small special-purpose agents? Here we describe a simple simulation pointing towards the second possibility, linking together two different VA's by a third VA.

Linked VA's and emergent attention

We began with two copies of the VA network discussed above, complete with reactive inhibition and distributed conflict monitoring. In every way these two copies of VA were identical, with one exception -- their sensors detected different kinds of objects. In our previous simulations, target color was not an issue, but here we introduced targets in two colors, red and green. One copy of VA -- the green VA -- had sensors that could only detect green targets; red targets were effectively invisible and would not activate the sensors in any way. The other VA copy -- the red VA -- could only detect red targets, and not green ones.

We will call the red and green copies of VA the *component* VAs.

The two component VAs were linked to a single shared set of motor units (see Figure 3). The sensor arrays of the component VAs were exactly aligned with each other, so that both component VAs were always looking at the identical region of space. So now consider what happens if we drop two green targets to this combination of VAs. The green targets are invisible to the red VA, and so the red agent provides no task-relevant activation to the motor units. The green VA would therefore have sole control of the motor system, and would proceed to catch the two targets, on its own, as it had already been evolved to do. Of course, the converse would be true if two red targets were dropped. For same-color targets, then, there is no cognitive conflict between the component VAs.

But now consider the case in which a red and a green target are dropped. Now each of the component VAs will attempt to catch the target it can see. In this case of mixed-color targets, the two VAs are in a state of cognitive conflict, one pulling towards red and the other towards the green target. Because there is no way to regulate the conflict between agents, we would expect interference from T2 on T1 catch in these conditions. Thus while each of the component VAs is a sophisticated information processor with memory and reactive inhibition, by linking them to the same set of motors we have turned the ensemble into something which might show little or no abilities for selective action.

To resolve this conflict, we added a third “agent”, that linked each of the hidden units in one component agent to each of the hidden units in the other. This third agent consisted only of this weight matrix. We refer to the collection of agents (the two component VAs and the third “conflict monitoring” agent) as the linked visual agent, or LVA. The weights and network parameters of the component VAs were fixed, and we evolved only the weights of the conflict monitor.

This small experiment was successful. LVA was able to process the mixed-color targets with high accuracy, a task which none of its components alone would be capable of. T1 and T2 catch accuracy were over 99%. Some results of this simulation are shown in Figure 4, which shows hesitation times for 500 novel same-color trials and 500 equivalent mixed-color trials. Recall that for same-color trials, there is no necessary conflict between the component VAs, while for mixed-color trials, there will be conflict. LVA showed an overall cost for reallocating resources from T1 to T2, a pattern very similar to VA. After catching T1, LVA hesitated for a signficant period before moving again to catch T2. Figure 4 illustrates hesitation to respond to T2 as a function of same vs mixed-color targets, and whether T2 was in-view or OOV after T1 catch. First let's consider LVA performance on same-color trials (ie two red or two green targets). As can be seen in Figure 4, LVA's performance on same-color targets was almost identical to the component VAs. In particular, hesitation was longer when T2 is in-view compared to OOV, $F(1,498)=137.8$, $p<0.0005$, suggesting that inhibition is proportional to target salience. This result is expected, and simply shows that the addition of the third agent -- the "conflict monitor" -- did not fundamentally changed the operation of the component VAs. The component VAs still used reactive inhibition to selectively respond to T1, and so were slower to reallocate processing towards salient T2s.

It is interesting to see a similar pattern even for the mixed-color trials. Again, when reallocating from T1 to T2 there was a significant hesitation, and this hesitation was reduced for the less salient OOV targets, $F(1,498)=78.5$, $p<0.0005$. Here then we also see evidence for reactive inhibition and increased inhibition for salient T2s. In this way, selective attention and action in the LVA is coordinated using mechanisms very similar to those in the component VAs.

There are several other interesting aspects to these results. First, while hesitation for LVA with same-color targets was very similar to those of the component VAs, hesitation was in fact slightly reduced for LVA, F(1,498)=7.52, p=0.006. Evidently LVA was able to reallocate more efficiently than a component VA. Recall that LVA consists of the unaltered component VAs plus the third "conflict monitor" agent, consisting of weights between the hidden layers of the component VAs. The better performance for LVA must then be due to the use of units in the "other" component VA, mediated by the conflict monitor weights. That is, on a trial with two green targets, LVA could reallocate more effectively than the component green VA could on its own. This suggests that LVA was using units in the red component VA to assist performance on trials with two green targets.

Further analysis suggested that this benefit was due primarily to more efficient processing of OOV T2s. The benefit for OOV over in-view T2s was larger for LVA than the component VA, as evidenced by the two-way interaction of Agent type (LVA or VA) and Visibility (in-view or OOV), F(1,498)=9.62, p=0.002. Hesitation for the in-view T2s was very similar (36.4 to 35.3 timeslices for VA and LVA respectively), with a larger difference for OOV T2s (28.3 to 24.8 timeslices). Although clearly this effect must be mediated by the conflict-monitor weights, additional analysis would be needed to determine how the "other" component VA facilitates reallocation to an OOV T2.

Another interesting result evident in Figure 4, is that reallocation was more efficient in the mixed than same color case, F(1,98)=320, p<0.0005. That is, it appears that cognitive conflict produced by the two targets was greater when both targets were loading on the same component VA. This suggests that T2 attracted less inhibition in the mixed than the same color case. This is to be expected since targets in the mixed case weren't competing for the resources of the same component VA. Less competition means less need for inhibition, and faster subsequent reallocation.

Finally, we have described the third VA here as the “conflict monitor”. We can be sure that this conflict monitor was responsible for the coordination of component VAs in the mixed-color case, since it is only these weights which allow the component VAs to communicate. Further analysis of these connections would therefore be very interesting, for example with respect to the specific proposals of Botvinick et al (2001). Perhaps these connections monitor the level of conflict between activation of motor units by the component VAs, or perhaps they use some other method for conflict evaluation and regulation.

Scaling up: More agents or Bigger agents?

The minimally cognitive agent was intended as a research tool, rather than a research hypothesis. We have seen the value of this tool in our previous studies, which show how understanding a minimally cognitive agent can help understand human performance. However, we would now like to undertake some wild speculation. We have seen how even a very small number of very simple processing units can exhibit sophisticated behavior. What if we do take seriously the idea of a small agent as a research hypothesis, concerning the basis of a cognitive architecture?

The idea of intelligent behaviour emerging from large collections of simple processors is an old one. Godfrey Thompson argued against the idea of Spearman’s general factor (g), suggesting instead a collection of many specific mental abilities, or “bonds”. The pros and cons of attempting complex cognition through a Minsky-style “Society of Mind” (Minsky, 1986) , or collection of many small agents, are reasonably well understood in computer science and robotics (e.g., subsumption architecture, Brooks, 1986).

However, the question of more or of more powerful agents is not often asked in neuroimaging and neuropsychology. Currently in psychology and neuroscience there is overwhelming agreement about the structure of the human brain's cognitive architecture, at least if we move to a high enough level of abstraction: the brain consists of a reasonably limited number of sophisticated, somewhat specialized processing systems. We'll come back to what is meant by a "reasonably limited number" later. In any case, the underlying assumption is that this limited number of cognitive modules are organized so that even simple tasks activate very large numbers of neurons. For example, it is typical to see reference to a small number of broad and powerful systems when interpreting brain imaging results, such as attention, memory, executive control, and so on.

We suggest the general assumptions of the standard view (we'll call it the "monolithic view") play out something as follows. What do we as psychologists and cognitive neuroscientists study? We study cognitive functions, specified at a fairly abstract level, for example, attention, or executive control. By way of comparison, most experimental psychologists would not say (admit?) that they study a specific task. Any single task, however well studied, is meant to give insights into more general cognitive mechanisms. The operating assumption is that the results from different tasks will converge to give a picture of these high level mechanisms, abstracted away from specifics of stimuli and specific presentation parameters. For example, visual search across space and across time might be assumed, or perhaps hypothesized is more accurate, to tap into a common set of attentional processes.

Similarly, it is assumed or hypothesized that there are component brain systems recruited for these different functions. A complex task like visual search might involve multiple such brain systems, including for example executive and visual memory systems. Each of these systems is meant to participate in many tasks.

So now let’s return to the issue of what is a “reasonably limited number” of component cognitive systems? It would be pointless to give a specific number, or even a specific order of magnitude. The important assumption of the monolithic view is surely that the number of component cognitive systems is much smaller than the number of tasks a person is capable of performing.

There are reasons to believe the monolithic view is correct. One is that brain imaging studies show considerable overlap in the brain areas activated during related tasks. For example, dorsolateral prefrontal cortex is active in a wide variety of demanding cognitive tasks (Duncan & Owen, 2000). One idea is that this area of cortex is analogous to a general-purpose computing facility (Duncan, 2001). Many neurons in the frontal lobe have firing properties that vary with task demands (e.g., Rao, Rainer, & Miller, 1997). Further support of the monolithic view comes from brain lesion studies. For example, damage to the tempero-parietal junction of the right hemisphere frequently results in unilateral spatial neglect, in which the left side of space is ignored. This is consistent with the possibility that a "spatial module" has been damaged, a module that provides input into many diverse kinds of spatial tasks (Halligan, Fink, Marshall, & Vallar, 2003).

At this stage, however, these lines of evidence provide a compelling argument for the monolithic view. Of course, one thing that makes brain imaging interesting is that differences as well as similarities can be found even between even highly related tasks. Likewise, the effects of brain lesion are frequently very hard to predict and there can be intriguing and surprising patterns of preserved and lost function. For example, in neglect there are frequent double dissociations between tasks like cancellation and line bisection (Ferber & Karnath, 2001).

More interesting is the fact that a single neuron might encode different kinds of information depending upon the current task, as in the Rao et al (1997) study. This kind of adaptibility in coding is necessary for a programmable, monolithic structure. However, this adaptibility is also perfectly consistent with a Society of Mind, and Minsky (1986) includes the notion of agents carrying references to other agents, in a task-dependent manner (e.g. “frames”).

Also, most psychologists will be familiar with the fractionation that occurs with increased study of a specific function. Take visual-object recognition for example. Current thinking seems opposed to the idea of a general-purpose visual recognition system. Instead, there appear to be multiple, more specific sub-systems for dealing with specific kinds of recognition problems. For example, there appear to be specialized face identification systems, areas for places, body parts, viewpoint-invariant and viewpoint specific recognition, holistic versus componential recognition, and so on. Likewise the concept of memory has fractionated into a variety of different forms of memory: short-term, long-term, epsisodic, auto-biographical, visual working memory, acoustic loop, and so on. A general trend seems to be that increased study of a cognitive function leads to evidence of more and more diverse forms of that function, rather than to grand unified theories.

In any case, our intention here is not to present a knock-down argument against a monolithic view, but to raise awareness about an alternative possibility, and speculate about how results with VA might inform a view more like the Society of Mind. Again, it is worth considering the orders of complexity separating VA and the human brain. VA: 17 simple "interate and fire" units, Human brain: on the order of 100,000,000,000 neurons, each a sophisticated information processing device. In principle, it appears that the human brain would have capacity for something on the order of billions of networks of VA's complexity. In such a collection, of very large numbers of agents, the organizing principle would small assemblies of units, each assembly capable of performing a very specific function. Everyday tasks, like talking on the phone, would require large numbers of agents, for parsing the words being spoken, anticipating the flows of conversation, activating appropriate schemas, modelling the speaker's mental state, planning and executing the next speech act, keeping the phone in the correct position, maintaining body posture, monitoring concurrent stimulation, and so on. Each of these small agents would possess its own selective processing system as needed, its own task-relevant memory system, task-specific object-recognition systems, and so on. Such a brain architecture would imply a hyper-distributed attention system, in which targets of action emerge from the combined processing many simple agents, perhaps on the order of thousands to millions.

There are some well known difficulties with the Society of Mind view. Perhaps most worrying is related to action selection. How is it determined which of perhaps billions of agents will control behaviour? What is to prevent an incoherent paralysis resulting from the activation of many conflicting agents? There are no easy answers, but certainly in other domains it is evident that a collection of many "agents" can produce purposeful behaviour. For example, it is not clear how cells maintain their specialization and organization to produce coherent structure and function at the level of the organism, but somehow they do. We do not want to trivialize these issues. Instead we want to suggest that the cognitive power of small assemblies, as demonstrated by VA and many other examples, may be well appreciated in the computer science community (e.g., Singh, 2003), but perhaps not in other domains of cognitive science. A cognitive architecture described by massive numbers of these assemblies may be a reasonable hypothesis for psychologists and neuroscientists.

References

Andersen, B. B., Korbo, L., & Pakkenberg, B. (1992). A quantitative study of the human cerebellum with unbiased stereological techniques. *Journal of Comparative Neurology*, *326*(4), 549-560.

Beer, R. D. (1996). *Toward the evolution of dynamical neural networks for minimally cognitive behavior*. Paper presented at the From animals to animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior, Cambridge, MA.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624-652.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Science*, *8*(12), 539-546.

Brooks, R. (1986). A robust layered control system for a mobile robot. *Robotics and Automation, IEEE Journal of [legacy, pre-1988]*, *2*(1), 14-23.

Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Science*, *3*(9), 345-351.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*(3), 332-361.

Critchley, H. D., Mathias, C. J., Josephs, O., O'Doherty, J., Zanini, S., Dewar, B. K., et al. (2003). Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. *Brain*, *126*(Pt 10), 2139-2152.

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, *92*(Pt 1), 53-78.

Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex.

Nature Reviews Neuroscience, *2*(11), 820-829.

Duncan, J. & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*(10), 475-483.

Fellows, L. K. & Farah, M. J. (2005). Is anterior cingulate cortex necessary for cognitive control? *Brain*, *128*(Pt 4), 788-796.

Ferber, S. & Karnath, H. O. (2001). How to Assess Spatial Neglect-Line Bisection or Cancellation Tasks? *Journal of Clinical and Experimental Neuropsychology*, *23*(5), 599-607.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Graham, B. P. & van Ooyen, A. (2004). Transport limited effects in a model of dendritic branching. *J Theor Biol*, *230*(3), 421-432.

Grison, S. & Strayer, D. L. (2001). Negative priming and perceptual fluency: more than what meets the eye. *Perception & Psychophysics*, *63*(6), 1063-1071.

Halligan, P. W., Fink, G. R., Marshall, J. C., & Vallar, G. (2003). Spatial cognition: evidence from visual neglect. *Trends in Cognitive Science*, *7*(3), 125-133.

Houghton, G. & Tipper, S. P. (1994). A model of inhibitory mechanisms in selective attention. In D. Dagenbach & T. Carr (Eds.), *Inhibitory mechanisms in attention, memory, and language.* San Diego, CA: Academic Press.

Koch, C. (1999). *Biophysics of Computation: Information Processing in Single Neurons.* Oxford University Press.

Mari-Beffa, P., Estevez, A. F., & Danziger, S. (2000). Stroop interference and negative priming: problems with inferences from null results. *Psychon Bull Rev*, *7*(3), 499-503.

Minsky, M. (1986). *The society of mind.* Simon & Schuster, Inc. New York, NY, USA.

Mulert, C., Menzinger, E., Leicht, G., Pogarell, O., & Hegerl, U. (2005). Evidence for a close relationship between conscious effort and anterior cingulate cortex activity. *Int J Psychophysiol*, *56*(1), 65-80.

Pakkenberg, B. & Gundersen, H. J. (1997). Neocortical neuron number in humans: effect of sex and age. *Journal of Comparative Neurology*, *384*(2), 312-320.

Rao, S. C., Rainer, G., & Miller, E. K. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, *276*(5313), 821-824.

Schneider, W. X. (1995). VAM: A neuro-cognitive model for visual attention control of segmentation, object recognition, and space-based motor action. *Visual Cognition*, *2*, 331-375.

Singh, P. (2003). Examining the Society of Mind. *Computing and Informatics*, *22*, 521-543.

Slocum, A. C., Downey, D. C., & Beer, R. D. (2000). *Further experiments in the evolution of minimally cognitive behavior: From perceiving affordances to selective attention*. Paper presented at the From animals to animats: Sixth International Conference on Simulation of Adaptive Behavior, Cambridge, MA.

Ward, R. (1999). Interactions between perception and action systems: A model for selective action. In G. W. Humphreys, J. Duncan, & A. Treisman (Eds.), *Attention, space and action: Studies in cognitive neuroscience.* Oxford: Oxford University Press.

Ward, R. & Ward, R. (in press). Selective attention and control of action: Comparative psychology of an artificial, evolved agent and people. *Journal of Experimental Psychology: Human Perception and Performance*.

Ward, R. & Ward, R. (2006). Cognitive conflict without explicit conflict-monitoring in a dynamical agent. *Neural Networks*, *19*(9), 1430-1436.

Figure captions

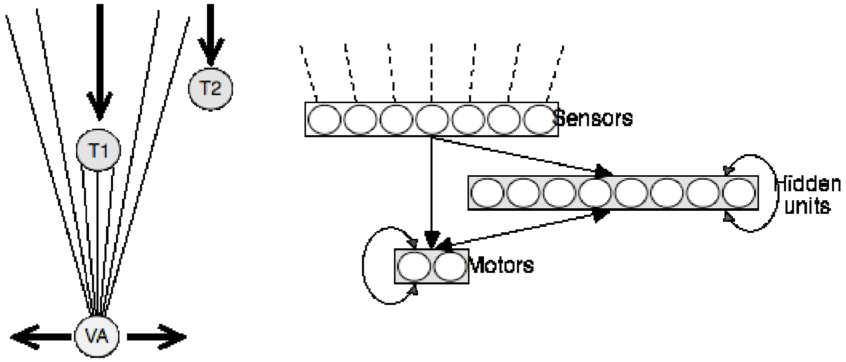
Figure 1. The visual agent (VA), originally developed by Slocum et al (2000). The left-hand side of the figure shows a schematic of the agent and environment. The agent is evolved to catch two targets (T1 and T2), falling from the top of the environment. VA has a left and a right motor, enabling it to move along the bottom of the environment. It is evolved by genetic algorithm to catch the targets as close to center as possible. The rays emanating from VA represent proximity sensors. The right-hand side of the figure shows the continuous-time recurrent neural network (CTRNN) which controls VA. The seven input units each pass on activation from an associated proximity sensor. The activation of the input units reflects where the proximity sensor is intersected. Each input unit is connected to the two motor units and to each of the eight hidden units. Each of the hidden units is receives connections from every unit in the network (including itself), and sends connections to every hidden and motor unit. The two motor units receive inputs from the input and hidden layer. These units also send connections back to the hidden layer, and both are self-connected and connected to each other.

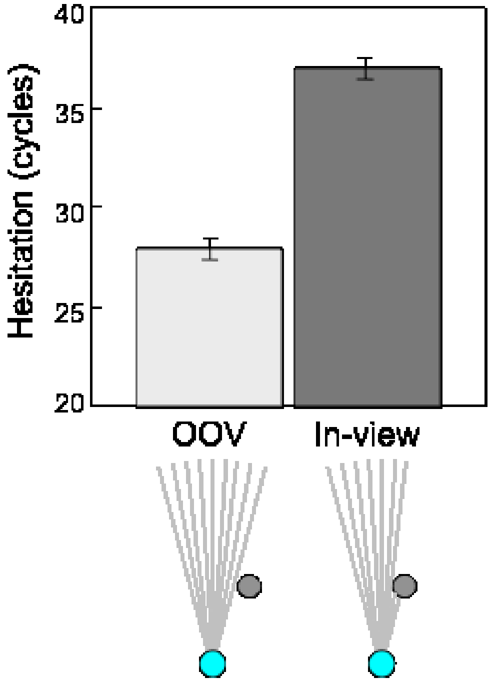
Figure 2. Hesitation after T1 catch. After catching the first target (T1), VA hesitates before moving again to catch the second target (T2). Hesitation is a function of T2 salience, and in particular whether T2 is in-view or out-of-view (OOV) following T1 catch. On in-view trials, T2 is within the sensor array following T1 catch; but not on OOV trials (as indicated by the icons below the graph). VA is faster to respond to T2 when it is OOV compared to in-view. This is consistent with a reactive inhibition account, as described in the text.

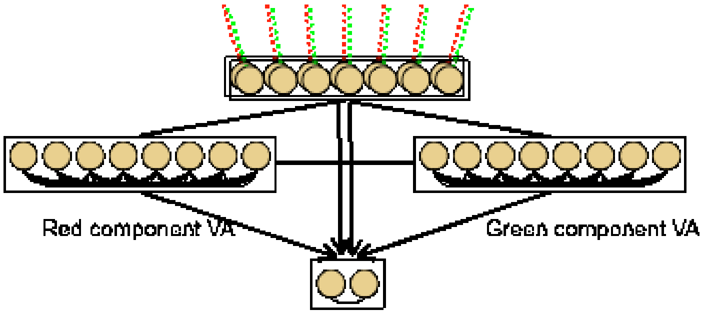
Figure 3. The Linked Visual Agent (LVA). LVA combines three agents. There are two *component* agents which have identical network structure and parameters as VA: the red

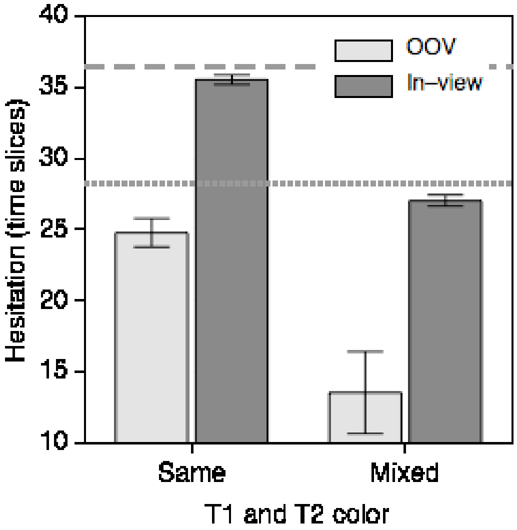
and green VA. Each of the component agents can only see targets in one color. The input and hidden layers of each component project to a shared motor layer. Thus the component VAs can be considered to compete for control of motor activity. Note that the proximity sensors of the red and green component VAs are exactly aligned so that they receive input from the identical region of space. Communication between the agents is added by a third "conflict monitor" agent, consisting solely of weights between the hidden layers of the two component agents.

Figure 4. Hesitation in LVA. Hesitation after T1 catch is shown for Same-color (T1 and T2 are both red or both green), and Mixed-color (one red target, one green target) trials. For comparison, hesitation of VA is shown for OOV (dotted line) and in-view (dashed line) trials.


T2
T1
VA
Sensors
Hidden units
Motors


Hesitation (cycles)
40
35
30
25
20
OOV
In-view


Red component VA
Green component VA


40
35
30
25
20
15
10
Hesitation (time slices)
OOV
In–view
Same
Mixed
T1 and T2 color