In Proceedings of the Sixth International Conference on Cognitive Modeling, 231-236. Mahwah, NJ: Lawrence Earlbaum.

Modelling performance in the Sustained Attention to Response Task

David Peebles (D.Peebles@hud.ac.uk)
Department of Behavioural Sciences, University of Huddersfield,
Queensgate, Huddersfield, HD1 3DH, UK

Daniel Bothell (db30@andrew.cmu.edu)
Psychology Department, Carnegie Mellon University,
Pittsburgh, PA 15213, USA

Abstract

We present a computational model of human performance on the Sustained Attention to Response Task, a computer-based task in which people must withhold responses to infrequent and unpredictable stimuli during a period of rapid and rhythmic responding to frequent stimuli. The model, formulated within the ACT-R cognitive architecture, accounts for human performance in terms of two competing strategies and the dynamic modification of priorities given competing task demands to minimise both response time and error. The model suggests that such strategic factors may be responsible for the observed speed-accuracy trade-off rather than the alternative proposal based on sustained attention.

Introduction

The Sustained Attention to Response Task (SART; Robertson, Manley, Andrade, Baddeley, and Yiend, 1997) is a computer-based task designed to measure a person's ability to withhold responses to infrequent and unpredictable stimuli during a period of rapid and rhythmic responding to frequent stimuli. As its name implies, the creators of the task interpret performance as reflecting the ability to sustain attention, which they define as "the ability to self-sustain mindful, conscious processing of stimuli whose repetitive, non-arousing qualities would otherwise lead to habituation and distraction to other stimuli" (Robertson et al., 1997, p. 747).

In the standard version of the task, participants are presented with the digits 1 to 9 in random order at a rate of one every 1.15 s. Each digit is presented for 250 ms followed by a 900 ms mask and participants are required to respond to each digit as rapidly as possible by clicking the mouse, apart from when they see the number 3 when they must withhold the response. The task consists of a total of 225 trials (25 of each of the 9 digits) and lasts approximately 4.3 min. Participants are instructed to respond as quickly as possible while making as few errors (i.e. incorrectly responding to a 3) as possible.

The central idea behind this paradigm is that the continuous performance over 225 trials together with the long and unpredictable intervals between targets encourages the development of an automatic response to non-target 'go' trials and that vigilant monitoring is required to withhold this response on the infrequent 'no-go' target trials. This contrasts with previous perceptual detection paradigms used to test sustained attention which have typically required participants to respond to the infrequent target (e.g., Loken, Thornton, Otto, & Long, 1995; Parasuraman, Mutter, & Molloy, 1991; Whyte, Polansky, Fleming, Coslett, & Calvallucci, 1995). Robertson et al. (1997) argue that in these paradigms, it is responses to the target stimuli that can become automatised which tends to increase the level of performance and makes the detection of attention lapses harder.

The main focus of research with the SART has been to investigate the performance of patients with traumatic brain injury (TBI) affecting the frontal lobes, a region previously associated with the ability to sustain attention (e.g., Rueckert & Grafman, 1996; Wilkins, Shallice & McCarthy, 1987). In several studies, Manly, Robertson, and their colleagues have shown that performance on the SART is diminished following TBI to the frontal lobes and that SART performance correlates with other measures of sustained attention (e.g., Manly, Robertson, Galloway, & Hawkins, 1999; Manly, Davison, Heutink, Galloway, & Robertson, 2000; Manly et al., 2003; Robertson et al., 1997).

The relationship between response time and error in the SART

In several studies, Manly and Robertson have found that individuals' rates of responding were significantly predictive of the number of errors they made (Robertson et al., 1997; Manly et al., 1999). In particular, the mean RT for the four trials prior to an error of commission (i.e. incorrectly responding to a 3) was found to be significantly faster than that for the four trials prior to a correct response to a no-go trial. This suggests that performance may be determined to a large extent by an individual's particular emphasis when trying to satisfy the competing task instructions to minimise both response time and the number of erroneous responses.

To study the relationship between RT and error in more detail, Manly et al. (2000) pooled data from 109 neurologically healthy participants who had carried out the SART in a number of different studies. They found that the mean RT for responses to a go trial was 375 ms (*SD* 65) and that subjects made on average 6.36 (*SD* 4.36) errors of commission and 1.06 (*SD* 3.41) errors of omission (incorrectly withholding a response on a go trial). They also found a significant negative correlation ($r = -0.49$) between the mean RT for go trials and the number of errors of commission made, indicating that,

Table 1: Mean Number of Responses and Mean Response Time (RT in ms) for Trials Immediately Before and After Correct and Erroneous Responses to the Target Digit 3 in the SART (Manly et al., 2000).

	Pre-correct		Pre-error		Post-correct		Post-error	
	Response	RT	Response	RT	Response	RT	Response	RT
Mean	17.6	384	6.0	333	17.30	351	6.07	364
SD	4.34	68.4	4.03	63.7	4.32	60.6	3.95	95.17

as in the previously reported studies, participants who responded more rapidly were more likely to make more errors. This relationship is shown in Figure 1 which plots the number of errors of commission against mean RT to non-targets for each participant.

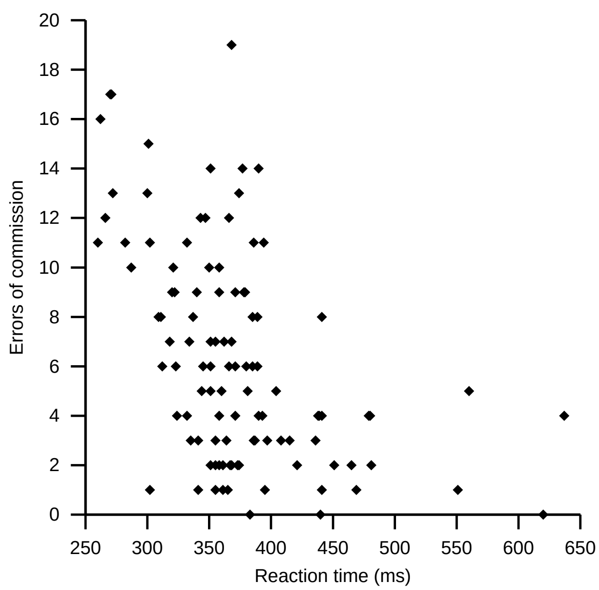

Figure 1: Plot of errors of commission against reaction time to non-targets in the SART (reproduced from Manly et al., 2000).

Manly et al. also analysed the within-subjects pattern of RTs relative to no-go trial responses. To do this they identified four categories of go trial: (a) those immediately prior to a correctly withheld response to the target digit (pre-correct); (b) those immediately prior to an incorrect response to the target digit (pre-error); (c) those immediately after a correctly withheld response to the target digit (post-correct); (d) those immediately after an incorrect response to the target digit (post-error). Their findings are summarised in Table 1. The mean RT for pre-error trials (333 ms) was significantly faster than that for pre-correct trials (384 ms). In addition, RTs for post-error trials (364 ms) were significantly slower than those for pre-error trials (333 ms), indicating that after participants had detected an error, they made a conscious decision to slow down. Finally, the data also revealed that the mean RT for post-correct trials was significantly faster than that for the pre-correct trials, indicating that in general, participants tended to speed up after correctly withholding a response to 3.

In an attempt to reduce the strength of the relationship between RT and error rates found in their study, Manly et al. (2000) produced a modified 'response-locked' version of the SART in which subjects were required to respond in time with an auditory signal presented 100 ms after stimulus onset. This modification was designed to reduce the variability in RTs both between and within subjects by setting a target RT close to that found in the previous study to be associated with correct responses to no-go trials (approximately 380 ms). Manly et al. argued that the modified task should be able to differentiate between the alternative accounts of the speed-accuracy trade-off in RTs. For example, if the main determinant of performance is the ability to sustain attention, then the number of errors of commission in the response-locked SART should be the same as in the standard version, although the amount of variation within an individual participant's RTs and also between participants RTs should be reduced. However, if participants' balancing of priorities in relation to the conflicting task demands is the primary determinant of performance, then setting the target RT to a value that has previously been associated with correct responses to no-go trials should reduce the number of errors of commission as well as the variability in RTs.

In an experiment in which 30 normal participants carried out both response-locked and standard versions of the SART, Manly et al. (2000) showed that the auditory signal significantly reduced the variability in individual participant's RTs and also between participants' RTs in the response-locked condition. They also found that the modification reduced the correlation between RT and error in the response-locked condition ($r = -0.19$). The comparative results were somewhat less conclusive. In the previous study, the proportion of erroneous responses to target trials was 25.44%. In the standard experiment condition, that proportion increased to 27.67% but in the response-locked condition it decreased to 22.57%. It seems, therefore, that introducing an auditory signal in the response-locked condition reduces the number of errors of commission as well as the variability in RTs, supporting the view that participants' balancing of priorities is a strong factor in performance. The difference between the two experiment conditions was found to be only approaching statistical significance, $F(1, 29) = 3.16$, $p = .09$. However the standard deviations of the errors for both the standard SART ($M = 8.30$, $SD = 5.11$) and response-locked SART ($M = 6.77$, $SD = 4.80$) are relatively large so it is possible that this issue may only be made less equivocal by a further study using a larger sample size.

A process account of the SART

The principal relationship between RT and error rate may, therefore, be characterised as a speed-accuracy trade-off, but as Manly et al. (2000) argue, this merely describes the relationship but does not provide an account of it. Manly et al. suggest that this pattern results from the waning of sustained attention that occurs during the course of the task. An alternative interpretation, however, is that observed performance is the result of *strategic* factors, in particular the attempt by subjects to find an appropriate strategy that allows them to satisfy the conflicting demands to minimise both RT and error. It is an open and interesting question, therefore, whether a strategy-based interpretation can account for performance in the SART. The purpose of the modelling work reported here is to address this question by producing and analysing a process model of the task.

The ACT-R cognitive architecture

ACT-R 5.0 (Anderson et al., submitted) is the most recent version of the ACT-R cognitive architecture (Anderson & Lebiere, 1998) that adds perceptual and motor modules to the existing cognitive module. The perceptual and motor modules provide ACT-R with rudimentary speech and audition capabilities, visual attention and processing mechanisms, and elements of motor control to simulate interaction with a computer keyboard and mouse.

ACT-R's cognitive system consists of a module for maintaining information about goals and two memories: a procedural memory in the form of a set of production rules and a declarative memory which is a network of chunks. The interface between the production rules and the other modules is through a set of *buffers*. The buffers hold information about such things as the current goal, the item of declarative knowledge that is currently available to the system, and the current state of the perceptual and motor modules. Each buffer may contain only one chunk of information, and each new request for information replaces the current contents of the buffer. Productions are rules of the form "IF <conditions> THEN <actions>," the conditions specifying a pattern of information that must be present in the buffers for the rule to apply and the action specifying the actions to be taken should this occur. The actions can be direct modifications of the buffers' contents or requests to other modules to change the buffers, for instance a request of the memory system to retrieve a chunk or a request to the visual system to encode a particular stimulus.

ACT-R is able to interact with a computer through its visual and motor modules. The visual system provides the model with the ability to detect changes in a computer's display and to attend to and encode the items that are there. The manual system allows the model to click a button on a simulated mouse or press a key on the keyboard. The timing parameters for those systems are built in and based on previous research. In particular the motor system is heavily influenced by the EPIC architecture (Kieras & Meyer, 1997).

Currently, ACT-R has one explicit attention mechanism, the primary purpose of which is to determine how activation from the current goal is used to retrieve knowledge in declarative memory. Processing in ACT-R is driven by the current goal (which is also represented as a chunk in declarative memory) and elements of the goal are viewed as the main focus of attention. ACT-R has a limited quantity of *source* activation (ACT-R's W parameter) which is shared equally between the elements in the goal and subsequently spread to associated chunks in declarative memory when retrieval requests are made. However, the SART is a very rapid, stimulus-driven respond-don't-respond paradigm which can be modelled in ACT-R with the assumption that participants are required to retrieve little, if anything, from declarative memory (the model assumes that the target number 3 is an element of the goal). As a consequence, this attention mechanism is not utilised in the model reported here.

A description of the model

An ACT-R model was constructed that was able to interact with the SART through the same interface as the human participants used—text presented on a computer screen and a mouse to enter the response. The model consists of 11 production rules, illustrated in Figure 2.

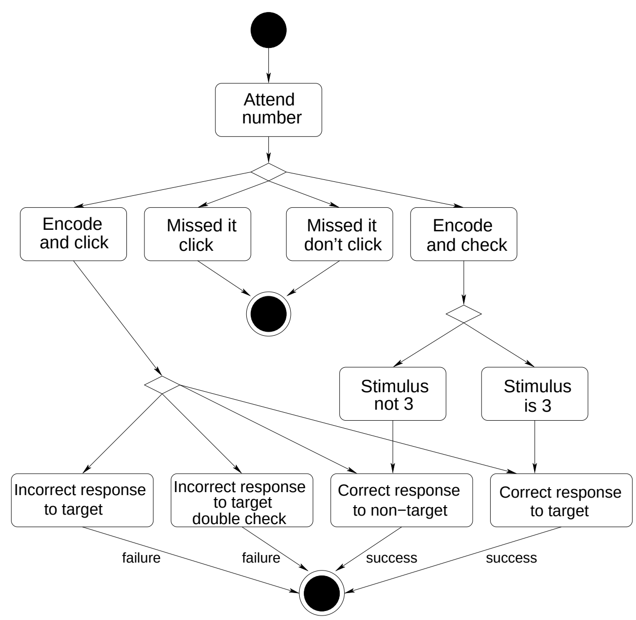

Figure 2: Flow of processing in the ACT-R SART model. Each state in the chart corresponds to one production rule in the model.

The model comprises two competing strategies, each represented by one production rule. The first is a faster, but less accurate method in which the model simply clicks the mouse after detecting the stimulus (encode-and-click). The second strategy is more accurate but slower. In this case, the model first checks the stimulus to ensure that it should click the mouse (encode-and-check), and then does so when appropriate.

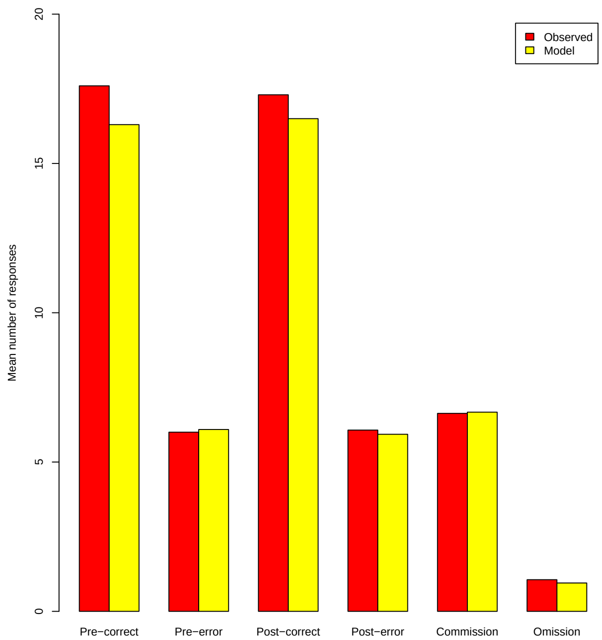

Figure 3: Observed and model mean number of responses for trials immediately before and after correct and erroneous responses to no-go trials in the SART and mean errors of commission and omission.

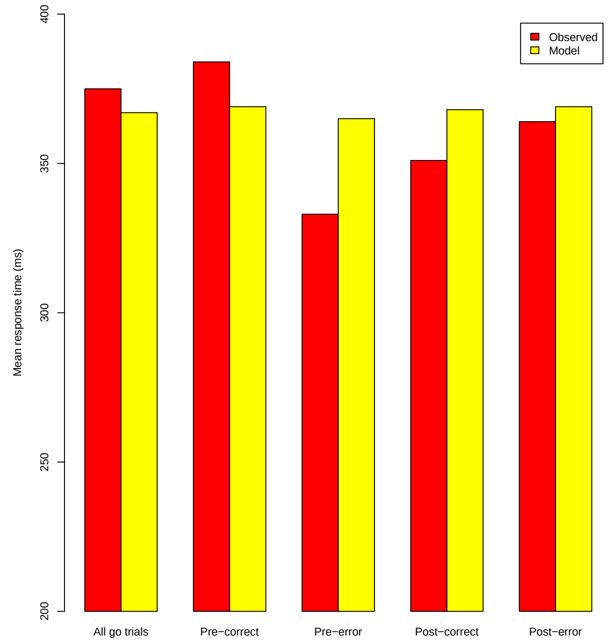

Figure 4: Observed and model mean response times for all go trials and trials immediately before and after correct and erroneous responses to no-go trials in the SART.

The choice between the two strategies is controlled by a quantity called *utility* which ACT-R uses to determine which production will fire next. The utility, U_i of a production i is defined as

$$U_i = P_i G - C_i + \sigma \tag{1}$$

where P_i is the probability of successfully achieving the goal if production i fires and reflects the history of successes and failures for the production, G is a global parameter that represents the cost (measured in time) of the current goal, and C_i is the measure of cost (also measured in units of time) associated with the use of production i until the goal has been achieved. When choosing from a number of candidate productions to fire, ACT-R's conflict resolution mechanism selects the production with the highest utility. To introduce an element of stochasticity to this process, however, a noise value, σ is added to each production's utility.

Initially, the utility values of the two productions that instantiate the strategies are equal. During the course of the task, the history of success and failure for each production, along with the time taken to produce those successes or failures, are used by ACT-R's utility learning mechanism to tune the P and C values of both productions. As a consequence, the model's preference for one strategy or the other changes from trial to trial.

In addition to the subsymbolic utility learning mechanism, the model also incorporates an explicit symbolic process for modifying its strategy. On each trial, after applying a strategy, the model verifies that it responded appropriately. In cases where the model verifies a correct response (the correct-response-to-target and correct-response-to-non-target productions) it simply waits for the next trial. If the model detects that it incorrectly clicked for a target, however, it can do one of two things; either just wait for the next trial (incorrect-response-to-target) or make an explicit choice to use the more accurate strategy on the next trial (incorrect-response-to-target-double-check). The production implementing this second option simply creates an additional element in the new goal to ensure that the encode-and-check production fires on the next trial, representing what is essentially an overt decision to be more careful in subsequent trials, (as evidenced by the significant increase in RTs for the post-error trials relative to pre-error trials observed by Manly et al., 2000).

There is one other aspect of the model can affect its performance. Like the EPIC architecture (Kieras & Meyer, 1997), completion times for ACT-R's perceptual-motor operations can be stochastic and one consequence of this is that the model sometimes fails to encode the stimulus during the 250 ms presentation period. On the relatively rare occasion that this happens, the model must decide whether to click the mouse (missed-it-click) or not (missed-it-don't-click). The structure of the SART implies that clicking the mouse is the more frequent response so a small preference for this option is implemented by giving the missed-it-click production a slightly higher initial utility value.

Applying the model

The ACT-R model was run 150 times to produce a data set from 150 simulated participants that could be compared with the RT and response data from the human subjects in the Manly et al. (2000) study. To fit the model to the data, two parameters which control the learning of production utilities were adjusted. One is the s parameter that adjusts the amount of the variance in the noise added to the calculations. The noise parameter was set to a low value (.01) to allow the model to be sensitive to the changes in the utility values of each production from trial to trial while still having some variability in its performance. The other is the G parameter representing the value of the goal in the utility calculation. The definition of G is based on the expected time to complete the goal and, as trial durations in the SART are very short (1.15 seconds) so is the expected goal completion time. Hence, the value of G was lowered from its default value of 20 seconds to 0.45 seconds to reflect this.

The results of the simulation are shown in Figures 3 and 4. The model provides a very close fit to the observed pattern of responses, accounting for 99% of the variance in the observed data (R^2 = .998, *RMSE* = .756). The model makes on average 6.67 errors of commission and 0.95 errors of omission (the latter occurring on occasion when the model fails to encode the stimulus). The RTs produced by the model are reasonably close to the observed data (R^2 = .665, *RMSE* = 19.933) but do not reproduce the pattern of variation between trial categories. However the mean RT for all go trials produced by the model (367 ms) is very close to the observed (375 ms).

The discrepancy between the observed and model RTs may be explicated somewhat by looking at the distribution of observed RTs in Figure 1, where it can be seen that the large majority of participants fall within a range of 340–400 ms. At either end of this range is a smaller group of participants who either respond very rapidly (between 250 and 340 ms) and make more errors or respond relatively slowly (between 400 and 650 ms) and make fewer errors. It is possible that the difference between the mean pre-correct and pre-error RTs is being magnified by the responses of these participants at the extreme ends of the range. The performance of these two groups of participants may reflect a strong preference for satisfying one of the task requirements over the other. In terms of the current model, this would be represented as a prior preference for a particular strategy, implemented by increasing the initial utility value for one of the two relevant productions.

To illustrate the range of behaviour covered by the model, Figure 5 plots the number of errors of commission against mean RT to non-targets for the 150 runs in the simulation. Although the range of RTs produced by the model is smaller than the observed, the model does produce a similar relationship between error and RT with faster means correlating with more errors. The correlation between mean RT on go trials and the number of errors of commission by the model is significant ($r = -0.788$, $p < .01$).

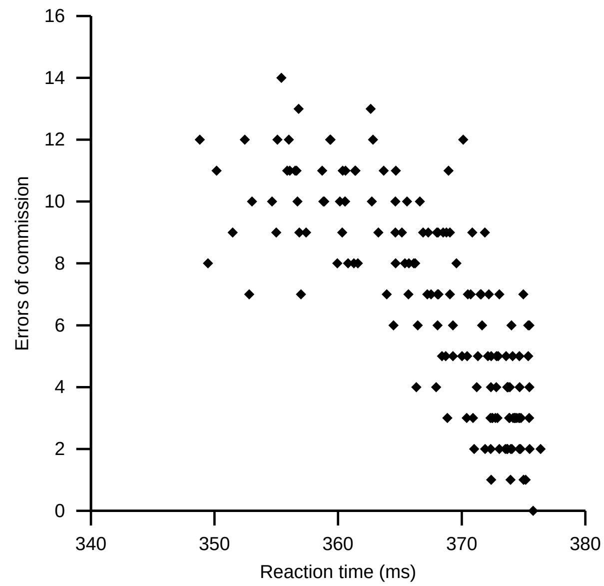

Figure 5: Plot of errors of commission against RT to non-targets for 150 runs of the ACT-R SART model.

Discussion

Although there is little doubt that the ability to sustain attention is an important factor affecting performance in many tasks, it remains an open question to what extent performance in the SART is determined by sustained attention or by the subject's attempt to satisfy conflicting task demands.

The model presented here questions the explanatory role of sustained attention in the SART by providing a plausible alternative account of performance in terms of strategy choice. The main process underlying this account is ACT-R's utility learning mechanism which increases the likelihood of a strategy being used if it has been successfully used before. As both strategies lead to a majority of successful trials, if the model starts to prefer one over the other, it will continue with that preference, but this choice of strategy will effect the number of errors and the mean RT produced by the model. The additional factor in the model is the explicit symbolic process for modifying its strategy after detecting that it has incorrectly clicked for a target. This has the effect of biasing the model in the direction of the encode-and-check strategy after an error of commission, resulting in the greater number of model runs producing slower RTs but fewer errors in the bottom right corner of Figure 5.

The RT and error data produced by the model provide a close fit to the observed data and, although the range of RTs currently captured by the model is limited to a range within the main section of the distribution, it does reproduce the error profile and the relationship between error and RT observed in Manly et al. (2000), which was the primary aim of the current effort. One proposal for future work under consideration is to use the current version of the model as a basis for analysing the performance of individuals at the extreme ends of the RT and error scales, possibly due to preferences in response selection (i.e. ACT-R's G parameter), small variations in strategy, or overall response speed (e.g., related to the speed parameters of ACT-R's motor module).

The phenomenon of speed-accuracy trade-off is commonly observed in human behaviour. In ACT-R, speed-accuracy trade-off is typically addressed by the conflict resolution mechanism for selecting the next production rule to fire—more specifically, the G parameter in Equation 1 (Anderson & Lebiere, 1998, p. 60). Changing the value of G affects the balance between speed and accuracy; a lower value of G places greater emphasis on the time cost of a production, C, resulting in a model that will be more likely to choose the rapid but less accurate option over the more successful, but slower strategy. Conversely, a higher G will increase the emphasis on successfully achieving the goal, even though this might take longer. This mechanism enables ACT-R models to account for a range of RT and error data along the trade-off spectrum. This model, however, is able to account for a range of behaviour in the SART with a fixed value of G. In addition, by incorporating an explicit symbolic level strategy to correct for detected errors in the model's responses, that range of behaviour was found to correspond to that of the human participants. The use of the utility learning mechanism to govern both the speed and accuracy of the model's behaviour is, to our knowledge, an original feature of this model.

It may be the case that Manly et al's (2000) first study can be completely accounted for in terms of the dynamic modification of priorities in relation to the conflicting demands of the task, without any reference to attention. However, an alternative conception of the model is also possible, one that views the differences in the amount of effort required to process a trial between the two strategies in terms of attention to the task. In this sense, the encode-and-click production represents the inattentive automatic response while the encode-and-check production embodies the additional 'attention' required to check the identity of the stimulus to successfully withhold the response to 3. This attentional interpretation of production rules has recently been adopted by other researchers attempting to model attention at the symbolic rule level (Wang, Fan, & Johnson, 2004), although it remains to be seen whether—and how far—this interpretation can be utilised in accounting for attention processes in executive control, particularly as production rules have traditionally been interpreted as representing unconscious, automatic, procedural knowledge. In terms of the new modular ACT-R 5.0 architecture, it is unclear whether a simple interpretation of attention at the production rule level will be adequate or whether an additional attention module and buffer may be necessary. It is clear, however, that the issue of how to understand and model attentional phenomena in cognitive architectures such as ACT-R is a pressing one that must be addressed.

Acknowledgements

The first author would like to thank Tom Manly for his kind help in providing information about the SART.

References

Anderson, J. R., & Lebiere, C. (1998), *The atomic components of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (submitted). An integrated theory of the mind. *Psychological Review*.

Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, *12*, 391–182.

Loken, W. L., Thornton, A. E., Otto, R. L., & Long, C. L. (1995). Sustained attention after closed head injury. *Neuropsychology*, *9*, 592–598.

Manly T, Robertson I. H., Galloway, M., Hawkins, K. (1999). The absent mind: Further investigations into sustained attention to response. *Neuropsychologia*, *37*, 661–670.

Manly, T., Davison, B., Heutink, J., Galloway, M., & Robertson, I. H. (2000). Not enough time or not enough attention?: Speed, error and self-maintained control in the Sustained Attention to Response Task (SART). *Clinical Neuropsychological Assessment*, *3*, 167–177.

Manly, T., Owen, A. M., McAvinue, L., Datta, A., Lewis, G. A., Scott, S. K., Rorden, C., Pickard, J. & Robertson, I. H. (2003). Enhancing the sensitivity of a sustained attention task to frontal damage. Convergent clinical and functional imaging evidence. *Neurocase*, *9*, 340–349.

Parasuraman, R., Mutter, S. A., & Molloy, R. (1991). Sustained attention following mild closed-head injury. *Journal of Clinical and Experimental Neuropsychology*, *13*, 789–811.

Robertson I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). Oops!: Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*, 747–758.

Rueckert, L., & Grafman, J. (1996). Sustained attention deficits in patients with right frontal lesions. *Neuropsychologia*, *34*, 953–963.

Wang, H., Fan, J., & Johnson, T. R. (2004). A symbolic model of human attentional networks. *Cognitive Systems Research*, *5*, 119–134.

Whyte, J., Polansky, M., Fleming, M., Coslett, H. B., & Calvallucci, C. (1995). Sustained arousal and attention after traumatic brain injury. *Neuropsychologia*, *33*, 797–813.

Wilkins, A. J., Shallice, T., & McCarthy, R. (1987). Frontal lesions and sustained attention. *Neuropsychologia*, *25*, 359–365.