# Title
Attention, Automaticity, and Priority Learning



Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 13(0)

Authors
Gupta, Prahlad
Schneider, Walter

Publication Date
1991

Peer reviewed

# Attention, Automaticity, and Priority Learning*

**Prahlad Gupta**
Department of Philosophy
Carnegie Mellon University
Pittsburgh, PA 15213
prahlad@cs.cmu.edu

**Walter Schneider**
Learning Research & Development Center
University of Pittsburgh
Pittsburgh, PA 15260
schneider@vms.cis.pitt.edu

## Abstract

It is widely held that there is a distinction between *attentive* and *automatic* cognitive processing. In research on attention using *visual search tasks*, the detection performance of human subjects in *consistent mapping* paradigms is generally regarded as indicating a shift, with practice, from serial, *attentional*, controlled processing to parallel, *automatic* processing, while detection performance in *varied mapping* paradigms is taken to indicate that processing remains under attentional control. This paper proposes a *priority learning* mechanism to model the effects of practice and the development of automaticity, in visual search tasks. A connectionist simulation model implements this learning algorithm. Five prominent features of visual search practice effects are simulated. These are: 1) in consistent mapping tasks, practice reduces processing time, particularly the slope of reaction times as a function of the number of comparisons; 2) in varied mapping tasks, there is no change in the slope of the reaction time function; 3) both the consistent and varied effects can occur concurrently; 4) reversing the target and distractor sets produces strong interference effects; and 5) the benefits of practice are a function of the degree of consistency.

## Introduction

Human performance changes dramatically as *practice* develops, leading to improved performance and a decrease in demands made on attentive resources. Many frameworks have discussed this transition (see [Shiffrin & Schneider 1977]). However, there is no generally accepted model of how performance changes with practice. Moreover, although theoretical frameworks have been proposed, they have generally not provided a detailed computational account of hypothesized processing.

In this paper, we present a computational model of practice effects in *visual search tasks* (see discussion below), which have played an important role in research on attentional processes. The term *attention* is generally used to indicate aspects of human cognitive processing that the subject can control, and that involve capacity or resource limitations. Attentional processing is taken to be a slow, serial activity, with the focus of attention limited to being one thing at a time. However, with practice in consistent tasks, *automatic processes* develop allowing parallel processing that is faster and not as limited by attentional resources[1].

The distinction between controlled and automatic processes has been the focus of much research in the field, particularly in visual search tasks. In such tasks, stimuli are presented visually to the subject, who is required to detect the presence of members of a set of target stimuli (the *memory set*). Nontarget stimuli are termed *distractors*. In *consistent mapping* paradigms, a target stimulus will never appear as a distractor on any trial – for example, the subject has to search the display for the memory set digits (e.g., the digits {8 4 5 2}) on all trials, with the remaining digits appearing as distractors. In *varied mapping* paradigms, targets on one trial may be distractors on another trial (e.g., the targets that the subject has to search for on a particular trial have been selected in the experiment by randomly sampling four of the ten digits, with the non-selected items being potential distractors; since a new sample is taken every trial, individual digits change roles as targets/distractors over trials).

A reliable finding from varied mapping studies is that reaction time increases roughly linearly with memory set size ([Kristofferson 1972a]). Consistent mapping studies have found that, with practice, search becomes much faster and memory set size has less impact on reaction times than in the varied mapping condition ([Neisser 1963, Neisser, Novick & Lazar 1963, Kristofferson 1972b, Schneider & Shiffrin 1977]).

These findings have led to the hypothesis that, in varied mapping, each item in the display is serially compared with items in the memory set, with each comparison taking on the order of 50 msec. In contrast, practice in the consistent mapping condition leads to a switch from serial, attentional, to parallel, automatic processing ([Shiffrin 1988]).

Our focus in this paper is on *priority learning*, which we propose as a mechanism that models the effects of practice, and the development of automaticity, in visual search tasks. We review specific experimental results that have been obtained in studies involving visual search tasks, and present simulations indicating how priority learning can provide an account of the observed practice and automaticity effects.

## An architecture for attentional effects in visual search tasks

This section describes briefly the larger architecture for attentional processing within which the priority learning model is embedded. The modular architecture (Figure 1) combines standard *connectionist components* such as connectionist units in a multi-layer organization (see [Rumelhart et al. 1986], for example) with *control elements* that modulate the flow of information between modules. The control involves a *gating unit* (Unit 1 in the figure) that provides a scalar multiplication of a module's output vector. The gating unit receives input from *priority units* within the module, and from an external *attentional control* that coordinates activity between modules. This architecture parallels certain aspects of neurophysiology, and is detailed elsewhere (see [Schneider & Detweiler 1987, Shedden & Schneider 1991, Schneider & Oliver, forthcoming]).

Figure 1 illustrates the overall model of a visual search task. "Visual" modules V1 and V2 represent areas of cortical visual processing, each corresponding to the small area of the visual field in which one stimulus appears. Each module consists of an *input layer* (labeled *I* in the figure) and an *output layer* (labeled *O*). The input layer projects via weighted connections to the output layer, which projects to a module at the next level of processing (Level 2). Each module also has a layer

*This work was supported in part by ONR contracts N00014-87K-0397 and N00014-86-K-0678 and ARI contract MDA903-89-K-0174 to the second author.

[1]See [Shiffrin 1988] for an excellent review of research on attention.

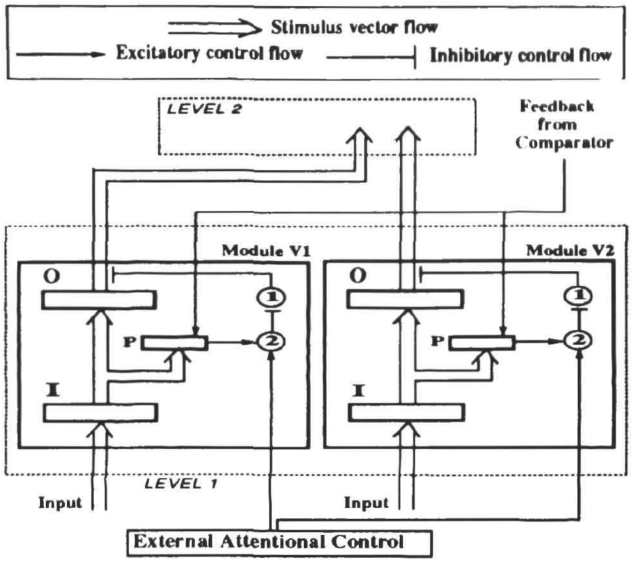


Figure 1: **Architecture of the model.** Stimuli from earlier stages of visual processing enter the input layers *I* of visual modules V1 and V2 and propagate forward to the output layers *O* (Level 1 of processing). The output layer vectors are sequentially transmitted (under external attentional control) to a comparator module (not shown), where they are serially compared with memory set items. Stimuli in the input layers of V1 and V2 also evoke a *priority level* activation over the layer of priority units *P*, which also receive feedback from the comparator module. The priority levels provide the basis for a transition from serial transmission and comparison of stimuli (*attentive* processing) to parallel, *automatic* processing (see text).

of priority units (labeled *P*), which receives input from the module's input layer.

Unit 1 within each module "gates" the forward propagation of activation from the output layer. When this gating is in effect, no vector *transmission* from the output layer of the module takes place, although the *activation* of the output layer is unchanged (e.g., as in the inhibition of the axon initial segment, but not the soma, of a neuron – see [Douglas & Martin 1990]). However, the external attentional control can (by exciting Unit 2, which inhibits Unit 1) inhibit Unit 1, which will release Unit 1's gating effect and lead to onward transmission of the vector in the output layer of that module. The attentional control thus provides for the sequential transmission of vectors from the visual modules, which models the shifting of attention from one area of the visual field to another, i.e., from one displayed stimulus to another, in the visual search task (see [Detweiler & Schneider, forthcoming]).

When the output layer vector of each module propagates forward in turn, it is compared serially with memory set items in a "comparator" module (not shown in Figure 1)[2]. If a visual stimulus matches a memory set item, then the output vector of that visual module also propagates forward to a "motor" module at Level 2 , initiating the motor sequence necessary to produce the detection response.

This paper describes the *priority learning* aspects of the model, whereby learning alters the activity of the priority units and allows automatic transmission of the output in the absence of attentional control.

## Priority learning

A sensory system that includes parallel input at early stages with serial processing at later stages would benefit from a filtering system that prioritizes inputs for limited serial processing. Early visual input processing is parallel ([Ericksen & Spencer 1969]). With a priority filtering scheme, a serial processing system can processes stimuli in their order of importance following the rating of stimuli based on a parallel low level interpretation of stimuli ([Norman 1969]). This allows the most important stimulus to be processed first. Humans appear to have such a priority filtering scheme ([Yantis & Johnson 1990]).

Our computational implementation of priority filtering is in the form of a connectionist network that associatively maps each input vector to a scalar priority. For example, in visual search, each letter would evoke a unique input vector of length 50 in a module (see Figure 1). Each input vector would associatively evoke a scalar value specifying the priority (on a scale of 0 through 10, represented by activation of the priority units) for that input vector. Input vectors evoking a high priority are transmitted to higher levels whereas low priority vectors are gated off, blocking the transmission. Many modules in the visual system can receive input simultaneously, with the respective input vectors each evoking a priority level, in parallel. If only one module contains a high priority target, it will "pop out" of the display ([Triesman 1988]). This occurs by having the priority units inhibit the gating units, thus allowing the output vector to be transmitted *automatically* to the next level of processing[3].

The priority units must be trained to associate input vectors to the appropriate priority levels. This is accomplished in the current model via feedback from the comparator stage. We assume that processing at the comparison stage results in higher output for more important stimuli ([Schneider & Detweiler 1987]). A scalar transform of this output signal is fed back to all modules at the earlier stage. This becomes the target level for the priority units of the transmitting module. Shortly after a vector is transmitted out of a module under External Attentional Control, the priority units are re-trained to approximate the feedback signal via associative learning.

The net result of priority learning is that if the stimulus in a visual module turns out to be a target, then the priority learning network is re-trained to evoke a *higher* priority than was currently evoked; if the stimulus turns out to be a distractor, then the priority learning network is re-trained to evoke a *lower* priority than was currently evoked.

The predicted effects of the learning of priorities are as follows. In a *consistent mapping* task over time, stimuli that enter modules V1 and V2 and correspond to targets will evoke increasingly high priorities in those modules, while stimuli

[2]This comparison can be performed by summing the transmitted vector with each memory set item into a layer of connectionist units. The sum of the squared activation of each unit in the receiving layer then provides a measure of the correlation, or *similarity* between the two vectors. If activation is above a criterion level, there is a *match* of the stimulus with a memory set item, otherwise no match. See [Shedden & Schneider 1991] for details.

[3]Figure 1 shows the excitatory connections from the priority units to Unit 2, which inhibits Unit 1. If the priority level evoked by a stimulus is above some threshold, Unit 2 can inhibit Unit 1 sufficiently to release its gating effect on the output layer, resulting in an "automatic" transmission from the visual module to the motor module. In this case, the motor response is triggered without any sequential scanning of modules under attentional control, or comparison of the stimulus in a module with memory set items. For simplicity, we have omitted discussion of control signals from the priority units to the external attentional control. These are described in [Schneider & Detweiler 1987], and provide a mechanism for external attentional control to entirely ignore a very low priority stimulus, i.e., to *fail* to initiate transmission from that module to the comparator.

corresponding to distractors will have increasingly low priorities. Once the priorities evoked by a target stimulus in modules V1 or V2 have crossed threshold level $\tau$, the presentation of that stimulus in either of those modules will lead to *automatic transmission* from the output layer of that module to the motor module, without sequential scanning of the modules, or serial comparison of their "contents" with memory set items. For distractor stimuli, low evoked priorities result in their being ignored by the attentional control. Thus, serial, attentional comparison is supplanted by parallel, automatic processing, resulting in reduced reaction times for the detection of targets, and in the independence of reaction time from memory set size.

In a *varied mapping* task, however, since a particular stimulus is sometimes a target and sometimes a distractor, there is no gradual increase/decrease of priorities for different sets of stimuli. Instead, the priority levels of stimuli will tend to fluctuate around the default priority level. That is, the evoked priority levels of presented stimuli do not cross threshold, and so there is no automatic over-ride of attentional processing, and no "ignoring" of stimuli. The production of the detection response has to continue to be through serial attentional comparison. Accordingly, there will be no decrease in reaction times, or in their linear relationship with memory set size[4].

## Implementation of priority learning

We ran simulations of the priority learning component of the architecture outlined above, using a single two-layer connectionist network incorporating the input layer and priority layer of one visual module shown in Figure 1. Our aim was to examine whether a priority learning scheme would provide a basis for the practice effects observed in consistent mapping and varied mapping paradigms.

Stimuli entering the network's input layer (*Input*, corresponding to layer $I$ in Figure 1) are vectors of length 50, with each element being in the range $\pm 1.0$; that is, the input layer consists of 50 units. The priority layer (*Priority*, corresponding to layer $P$ in Figure 1) consists of 10 units[5]. A stimulus in *Input* is transmitted in feed-forward fashion to *Priority*. Net input to each priority unit is computed as the weighted sum of inputs from all the input units. The *priority level* evoked by a stimulus is computed by a simple count of the number of priority units in *Priority* that have positive net input. Thus if all ten priority units have negative net input, the priority level is computed as 0; if $n$ priority units have positive net input, the priority level is $n$, with the highest possible priority being 10. Bias to the priority units is distributed linearly over the range $\pm 0.40$, so that five priority units have a positive bias, and five a negative bias. In the absence of input to the priority units from *Input*, therefore, the evoked priority level has the default value of 5.

When a stimulus is presented to the network, a priority vector is evoked at the priority layer. The network is then re-trained to evoke, for the current stimulus, either a higher priority level than currently evoked (if the current stimulus is designated a target), or a lower priority level than currently evoked (if the current stimulus is designated a distractor)[6]. This will be referred to as the *incrementing* and *decrementing* of priority, respectively. The simulations described here have employed the Widrow-Hoff learning algorithm ([Widrow & Hoff 1960]). One *epoch* consisted of presentation of the entire set of stimuli, with appropriate incrementing/decrementing of priority after presentation of each stimulus.

## Simulations of priority learning

**Simulation 1: Consistent mapping.** In consistent mapping studies with human subjects, the rate of search has been shown to become much faster with practice ([Neisser 1963]). Studies with well-practiced human subjects have exhibited reaction times that varied relatively little with memory set size ([Schneider & Shiffrin 1977, Experiment 2], [Neisser, Novick & Lazar 1963]).

In our simulations of consistent mapping, the training set for a module consisted of 16 random vectors, which were partitioned into two disjoint sets of 8 stimuli each, one set being designated *targets*, and the other *distractors*. During priority learning, the evoked priorities were consistently incremented or decremented for the two sets of stimuli respectively.

The results of the consistent mapping simulations are shown in Figure 2a. The slopes show the evoked mean priority level of targets and distractors, as a function of training (epochs). With increased epochs of training, target stimuli come to evoke increasingly high priorities, while distractor stimuli have increasingly low priorities. As discussed in the section on priority learning, a visual module in which a stimulus evokes a priority level of greater than some threshold $\tau$ will automatically transmit the vector in its output layer.

Figure 2b shows reaction times for targets in consistent mapping experiments with human subjects, who were trained over a period of 36 days ([Kristofferson 1972b]). Memory set sizes of one, two and four were used. For each set size, there was a decrease in reaction time over the training period, which is interpreted as being largely due to speeding up of non-attentional components such as the motor response component of the task, which is also the case in varied mapping tasks ([Shiffrin 1988, page 748]). Figure 2b also illustrates the non-linear set size functions characteristic of consistent mapping.

The priority levels of targets in our simulations of consistent mapping give rise to similar reaction times. On each epoch, search time $S$ for a target was simulated as follows. If the priority level evoked by the stimulus was greater than a threshold level $\tau$, then the stimulus would evoke automatic detection, taking "automatic response time" $A$. If the priority level was below threshold, then the target would have to be compared serially with memory set items, with each such comparison taking a constant "comparison time" $C$. We assumed non-terminating search, so that the number of comparisons required is equal to memory set size $m$. Total reaction time for a target was therefore calculated as (i) a base time $B$, representing non-attentional factors, *plus* (ii) search time $S$, which was either (a) automatic detection time $A$ (if the priority level was greater than $\tau$), or (b) controlled response time equal to comparison time $C$ times memory set size $m$ (if the priority level was less than $\tau$).

The time per comparison $C$ was taken to be 50 msec, and the time for an automatic response, 40 msec. We used priority threshold $\tau = 8.0$. We used a base reaction time $B$ of 290 msec, which decayed to about 230 msec over 200 epochs of training, simulating the speed-up of non-attentional components of reaction times[7]. Figure 2c shows the simulated reaction times,

[4] Atkinson & Juola ([Atkinson & Juola 1973]) have presented a similar model of recency learning, in which stimuli of intermediate familiarity require attentional scanning while novel stimuli do not.

[5] Although our simulations have used a *layer* of 10 priority units, it should be possible to achieve the same effects with a single priority unit.

[6] The higher or lower priority levels correspond to modulation of priority layer activations by feedback from the comparator module.

[7] These figures were derived from the human subject data by subtracting the time for one attentional comparison (50 msec) from total reaction time for memory set size one (a) at Days 1-6 (approximately 340 msec), and (b) at

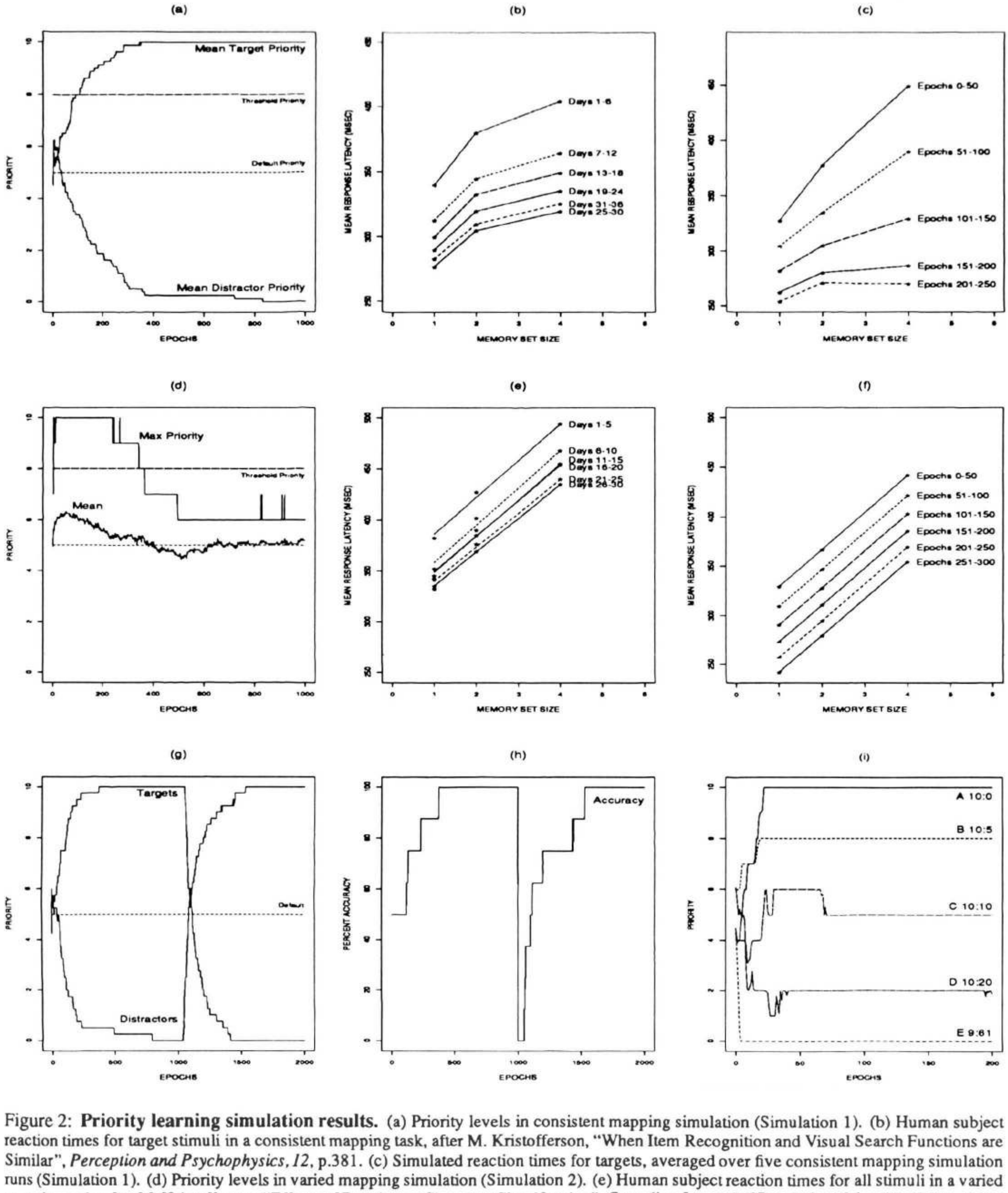


Figure 2: **Priority learning simulation results.** (a) Priority levels in consistent mapping simulation (Simulation 1). (b) Human subject reaction times for target stimuli in a consistent mapping task, after M. Kristofferson, "When Item Recognition and Visual Search Functions are Similar", *Perception and Psychophysics, 12*, p.381. (c) Simulated reaction times for targets, averaged over five consistent mapping simulation runs (Simulation 1). (d) Priority levels in varied mapping simulation (Simulation 2). (e) Human subject reaction times for all stimuli in a varied mapping task, after M. Kristofferson, "Effects of Practice on Character Classification", *Canadian Journal of Psychology 26*, p.57. (f) Simulated reaction times for all stimuli, in varied mapping simulation (Simulation 2). (g) Priority levels with reversal of targets/distractors in consistent mapping (Simulation 4). (h) Detection accuracy with reversal of targets/distractors in consistent mapping (Simulation 4). (i) Priority levels of stimuli with varying degrees of consistency (Simulation 5).

for memory set sizes of one, two and four. Levels of practice are shown in terms of average reaction time over blocks of 50 epochs, corresponding to the average reaction times over blocks of days in Figure 2b.

As in the human data (Figure 2b), the simulation results (Figure 2c) show a downward shift (with practice) of reaction times for each memory set size; this is the combined result of decreasing base response times as well as of developing automatic responses. Moreover, the reaction times for different set sizes at a particular level of practice exhibit the non-linearity expected in a consistent mapping task. Overall, the simulation results appear to model the human data quite well[8].

**Simulation 2: Varied mapping.** Varied mapping studies with human subjects have found that search time continues to be a linear function of memory set size, even after practice ([Kristofferson 1972a]). In our simulations of varied mapping, the training set for a module consisted of 16 random vectors. The priority level for each stimulus was incremented or decremented at random, modeling the effect of stimuli not being treated consistently either as targets or distractors.

Results from the varied mapping simulations are shown in Figure 2d, which plots the mean evoked priority of all stimuli. The mean priority levels fluctuate around the "default" priority level of 5, even with extended practice, as a result of the non-consistent treatment of stimuli. Note that in the simulation results shown, there is an initial random increase in mean priority; in other simulations, there was an initial decrease in priorities. But in all cases, the mean priority settled around the value of 5.

Human reaction times from a varied mapping task ([Kristofferson 1972a]) are shown in Figure 2e (all stimuli), which shows that detection times remain linear with memory set size, even with extended practice. The priority levels of stimuli in our varied mapping simulation give rise to similar reaction times, calculated in the same way as for consistent mapping (Figure 2f). As in the human subject data, practice does not change the linearity of set size functions. The reason for this is that, since few stimuli achieve threshold priority levels, serial comparison remains the only basis for target detection. Automatic detection does not develop, and therefore search functions remain linear with set size. Thus our simulation provides a good match with the observed varied mapping results[9].

**Simulation 3: Consistent & Varied mapping combined.** Schneider and Fisk conducted an experiment in which subjects were asked to carry out a consistent mapping task and a varied mapping task simultaneously ([Schneider & Fisk 1982a]). The findings were that both tasks could be accomplished simultaneously about as well as when each task was carried out alone.

We ran simulations in which 32 stimuli were partitioned into two sets of 16 stimuli each. One set of 16 stimuli represented a consistent mapping task: 8 of the stimuli were treated as targets, and the other 8 as distractors. The other set of 16 stimuli represented a varied mapping task, and the stimuli were treated randomly as targets and distractors. One epoch involved presentation of all 32 stimuli, and represented the simultaneous performance of both the consistent and varied mapping tasks on different sets of stimuli.

The priority levels of consistently mapped stimuli (which were consistently either targets or distractors) separated as in the consistent mapping task alone. The average evoked priority of the varied mapping stimuli remained close to the default priority level of 5, as in the varied mapping task alone. Thus the priority levels of both the consistent and varied mapping stimuli were similar to those in each of the tasks performed independently, indicating that practice effects in the combined consistent and varied mapping task can be modeled in terms of priority learning under the same assumptions as for the independent tasks, in conformity with the findings of the experiment described above.

**Simulation 4: Target/distractor reversal in consistent mapping.** Implicit in the priority learning scheme we have presented is the hypothesis that, in consistent mapping, with practice, target stimuli come to "attract attention" to themselves. Shiffrin and Schneider empirically verified the prediction that if the set relationships are reversed after the priorities have been established, subjects' performance deteriorates below that of novice subjects. For example, in initial practice, targets might be {A B C}, with {X Y Z} as distractors. After reversal, {X Y Z} would be the targets, and {A B C} the distractors. This reversal is expected to produce deterioration because the stimuli that were previously targets attract attention away from the previous distractors that become the new target stimuli.

In the experiments, subjects practiced a consistent mapping task ([Shiffrin & Schneider 1977, Experiment 1]). Once performance had stabilized (2100 trials), targets and distractors were interchanged. On reversal, detection accuracy dropped to below the level it had been at the beginning of the original consistent mapping task (approximately 50 percent), and 900 trials were needed for accuracy to return to the novice level with targets and distractors reversed. Subsequently, gradual re-learning occurred, and accuracy reached the previously established level.

We simulated this reversal in a consistent mapping task. A consistent mapping simulation was run with a set of target stimuli and a set of distractor stimuli for 1000 epochs. At the end of 1000 epochs, the targets and distractors were reversed, i.e., the priorities of the former distractors were now consistently incremented, and those of the former targets were now consistently decremented. The simulation was run in this reversed condition for a further 1000 epochs.

Figure 2g shows that the mean priority levels of the original targets/distractors increase/decrease over the first 1000 epochs of consistent mapping training. When targets and distractors are reversed at epoch 1001, the priority levels of former targets start dropping sharply, and those of former distractors start increasing sharply.

We computed an accuracy measure as follows: a given stimulus presentation was considered to have evoked an accurate response either if (i) the stimulus was a target, and evoked a priority of above threshold, or (ii) it was a distractor, and evoked a priority of below threshold. The threshold priority level used was 9.50. Figure 2h shows the variation of this

Days 31-36 (approximately 280 msec). This decrease in base time, from 290 msec to 230 msec, was fitted by an exponential function with constant 0.9985, as a first approximation.

[8] Our model predicts that reaction time functions will become perfectly flat at some point. This does occur in our simulations, by approx. 400 epochs of training.

[9] Note that the behavior of the *slopes* of simulated reaction times fit the human data quite well, which was the major focus of interest in this paper. By contrast, the changes in intercept (i.e., base reaction time) in simulations do not fit the human data well. This is because simulation of the base time practice effect did not include an asymptote effect and hence overestimated the decrease in base time. With humans, decreases in base time seem to asymptote at approximately 200 msec. Adding an additional parameter to the model would allow a good fit of both the constant slope and declining intercept effect in varied mapping.

accuracy measure. At the start of initial consistent mapping training, accuracy was approximately 50 percent, and this improved steadily, asymptoting by about 500 epochs of training. When targets and distractors were reversed at epoch 1001, accuracy dropped sharply, to below the level at the start of the simulation, and it required approximately 100 epochs of training to reach that level (50 percent). Beyond this point, re-learning occurred gradually, and by epoch 1500, accuracy had returned to the previous maximum. Thus the simulation models the human data fairly well.

**Simulation 5: Degree of consistency.** Schneider and Fisk found that consistency has a graded rather than an all-or-none effect. In one experiment, subjects searched for one letter (i.e., memory set size 1) in frames containing four letters ([Schneider & Fisk 1982b, Experiment 1]). Subjects were required to indicate the frame position in which the target had appeared, at the end of each trial. The degree of consistency of appearance of letters as targets or distractors varied. Thus one letter had the consistency ratio 10:0, meaning that it appeared as a target in 10 trials per block of trials, and never as a distractor. Three other letters had the consistency ratios 10:5, 10:10 and 10:20, respectively. Five other letters each had degree of consistency 9:61.

Throughout the experiment, detection accuracy was highest for the 10:0 letter, followed closely by the 10:5 condition. Detection accuracy was lowest for the 10:20 and 9:61 conditions, which did not differ substantially from each other. The 10:10 letter showed intermediate detection accuracy.

In our simulation, our training set consisted of five stimuli, which we designated A, B, C, D and E. One epoch consisted of the appearance of each of the letters A, B, C and D 10 times as a target, and 0, 5, 10 and 20 times, respectively, as a distractor; thus these letters corresponded to the consistency ratios 10:0, 10:5, 10:10 and 10:20, respectively. The E stimulus appeared 9 times as a target and 61 times as a distractor, corresponding to the consistency condition 9:61. The simulation involved presentation of this training set to a network representing the priority learning component of a single module. By approximately 70 epochs of training, the priority levels of the A, B, C, D and E stimuli had settled to 10, 8, 5, 2 and 0, respectively (Figure 2i).

Although we have not constructed simulations corresponding to the overall architecture that would be required for a simulation of the experiment described above, the priority learning results from the simulation with a single module suggest accuracy would be a graded function of the degree of consistency. The overall model would incorporate four modules V1, V2, V3 and V4, corresponding to the visual fields in which each of the four letters in a frame appear. The process of training would lead to each of the modules having the priority levels described above, which developed for a single module. A winner-take-all network between modules would lead to processing of stimuli in their order of priority, so that attention would be preferentially attracted to a higher-priority item even when it is not the target. Consequently, accuracy would be greater for higher-priority items.

## Conclusions

This paper has examined the viability of proposed "priority learning" mechanisms as the basis of a computational model of a number of the phenomena observed in visual search tasks performed by human subjects. The simulation results that have been presented are consistent with a fairly broad set of experimental findings. We therefore conclude that priority learning provides a fairly good account of practice and automaticity effects in visual search tasks.

## References


[Atkinson & Juola 1973] Atkinson, R. & Juola, J. 1973. Factors influencing speed and accuracy of word recognition. In S. Kornblum (ed.), *Attention and performance IV*. New York: Academic Press.

[Detweiler & Schneider, forthcoming] Detweiler, M. & Schneider, W. Modeling the acquisition of dual task skill in a connectionist/control architecture. In D. Damos (ed.), *Multiple-task performance: Selected Topics*. London:Taylor & Francis. Forthcoming.

[Douglas & Martin 1990] Douglas, R. & Martin, K. 1990. Control of neural output by inhibition at the axon initial segment. *Neural Computation* 2: 283-292.

[Ericksen & Spencer 1969] Ericksen, C. & Spencer, T. 1969. Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology Monograph* 79 (2, Part 2).

[Kristofferson 1972a] Kristofferson, M. 1972a. Effects on practice on character classification performance. *Canadian Journal of Psychology* 26: 54-60.

[Kristofferson 1972b] Kristofferson, M. 1972b. When item recognition and visual search functions are similar. *Perception & Psychophysics* 12: 379-384.

[Neisser 1963] Neisser, U. 1963. Decision time without reaction time: Experiments in visual scanning. *American Journal of Psychology*. New York: Appleton-Century-Crofts.

[Neisser, Novick & Lazar 1963] Neisser, U., Novick, R., & Lazar, R. 1963. Searching for ten targets simultaneously. *Perceptual and Motor Skills* 17: 955-961.

[Norman 1969] Norman, D. 1969. *Memory and attention: An introduction to human information processing*. New York: John Wiley.

[Rumelhart et al. 1986] Rumelhart, D., McClelland, J. & the PDP Research Group 1986. *Parallel Distributed Processing, Volume 1: Foundations*. Cambridge, Mass.: MIT Press.

[Schneider & Shiffrin 1977] Schneider, W. & Shiffrin, R. 1977. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review* 84: 1-66.

[Schneider & Fisk 1982a] Schneider, W. & Fisk, A. 1982a. Concurrent automatic and controlled visual search: Can processing occur without resource cost? *Journal of Experimental Psychology: Learning, Memory and Cognition* 8: 261-278.

[Schneider & Fisk 1982b] Schneider, W. & Fisk, A. 1982. Degree of consistent training: Improvements in search performance and automatic process development. *Perception & Psychophysics* 31: 160-168.

[Schneider & Detweiler 1987] Schneider, W. & Detweiler, M. 1987. A connectionist/control architecture for working memory. In G. Bower (ed.), *The Psychology of learning and motivation*, Vol. 21. Orlando, FL: Academic Press.

[Schneider & Oliver, forthcoming] Schneider, W. & Oliver, W. An instructable connectionist/control architecture: Using rule-based instructions to accomplish connectionist learning in a human time scale. In K. Van Lehn (ed.), *Architecture for intelligence*. Hillsdale, NJ: Lawrence Erlbaum. Forthcoming.

[Shedden & Schneider 1990] Shedden, J. & Schneider, W. 1990. A connectionist model of attentional enhancement and signal buffering. In Proceedings of the Twelfth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum.

[Shedden & Schneider 1991] Shedden, J. & Schneider, W. 1991. A connectionist simulation of attention and vector comparison: The need for serial processing in parallel hardware. In Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum.

[Shiffrin & Schneider 1977] Shiffrin, R. & Schneider, W. 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review* 84: 127-190.

[Shiffrin 1988] Shiffrin, R. 1988. Attention. In R. Atkinson, R. Herrnstein, G. Lindzey & R. Luce (eds.), *Steven's handbook of experimental psychology*. New York: John Wiley.

[Triesman 1988] Triesman, A. 1988. Features and Objects: The 14th Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology* 40A: 201-237.

[Widrow & Hoff 1960] Widrow, G. & Hoff, M. 1960. Adaptive switching circuits. In Institute of Radio Engineers, Western Electric Show and Convention, Convention Record, Part 4, 96-104.

[Yantis & Johnson 1990] Yantis, S. & Johnson, D. 1990. Mechanisms of attentional priority. *Journal of Experimental Psychology: Human Perception and Performance* 16: 812-825.