

A Theory of Saliency Change Dependent on the
 Relationship between Discrepancies on Successive Trials
 on which the Stimulus is Present

by Richard Sutton

ABSTRACT

All current theories of saliency change (Frey and Sears, 1978; Kirk, 1974; Mackintosh, 1975; Dickinson and Mackintosh, 1979; Moore and Stickney, 1980; Pearce and Hall, 1980) determine change in saliency based solely on events occurring on a single trial. However, a theoretical view of classical conditioning as prediction suggests that the relationship between the discrepancies on different trials on which a stimulus occurs should be used to change its saliency. This paper presents a theory of saliency change based on this idea and discusses its relationship to known experimental results and other theories of saliency change. Several experiments are proposed where the theory makes novel predictions subject to experimental test. A mathematical derivation of the main equation of the theory from high level assumptions is presented in the appendix.

1.0 The Proposed Saliency Change Rule

Following the form of the Rescorla-Wagner theory (Rescorla and Wagner, 1972), we write

$$\Delta V_A = \alpha_A [\lambda - V], \quad (1)$$

as the change in associative strength V , of stimulus A as a

function of its salience α_A , the intensity of the ucs λ , and the net associative strength term \bar{V} . Implicit in this equation is that it is only applied for those trials on which stimulus A occurs. Similarly, only those ^{stimuli} α_A present on the trial contribute to the total associative strength term \bar{V} . The simplest assumption is that \bar{V} is just a sum:

$$\bar{V} = \sum V_A, \text{ for all stimuli A present on the trial.}$$

\bar{V} is, in a sense, a prediction of what the reinforcement level λ will be. The difference between the predicted and the actual reinforcement level is usually called the discrepancy, D. In this case we have

$$D = \lambda - \bar{V} \quad (2)$$

All theories of salience, including the one proposed here, change the salience of a stimulus only on the trials on which the stimulus occurs. The theory being presented here is unlike other theories, however, in that the change in salience is dependent on the relationship between the results of more than one trial on which the stimulus occurs. To write the equation for the change in salience of a stimulus, we will need a notation to refer to the members of the entire sequence of trials on which the stimulus occurs. We will let trial T refer to the current trial, in which the stimulus must have occurred in order for its salience to be

changed, and trials T-1, T-2, T-3 etc, refer to the previous trials in which the stimulus occurred, trial T-1 being the most recent such trial. Using this notation we can then rewrite the Rescorla-Wagner rule (Equation 1) as

$$\Delta V_A^T = \alpha_A^T \cdot D^T$$

The theory of salience proposed here adds to this the following rule for changing the salience parameter α :

$$\alpha_A^T = \alpha_A^{T-1} + \theta_A \cdot D^T \cdot \left\{ D^{T-1} + (1-\alpha_A^{T-1}) \cdot D^{T-2} + (1-\alpha_A^{T-1})(1-\alpha_A^{T-2}) \cdot D^{T-3} + \dots \right\} \quad (3)$$

where θ_A is a small positive constant, and α_A can be assumed to typically lie between 0 and 1.0. In words, the change in salience of a stimulus is proportional to the product of the current discrepancy and a sum of the discrepancies of previous trials on which the stimulus occurred. These old discrepancies are weighted less and less heavily in the sum the farther back they appear in the sequence of previous trials in which the stimulus occurred. This decreased weighting of older terms is due to their multiplication by more terms of the form $(1-\alpha)$, which can be assumed to always be less than 1.0 in absolute value.

This salience change rule may seem somewhat complex. However, as briefly presented in the appendix, this rule can be

mathematically derived from the assumption of the Rescorla-Wagner rule and that salience α is changed towards minimizing the mean-square error between prediction \bar{V} and actual reinforcement λ . These assumptions are made more precise in the appendix where the derivation is presented. For now I only wish to point out that even though moderately complex, the exact form of this equation has been chosen based on a fairly small set of theoretically reasonable and interesting high level assumptions.

According to Equation 3, salience will increase if the current discrepancy is of the same sign, positive or negative, as the discrepancy on the most recent previous trials on which the stimulus occurred. Conversely, the salience will decrease if the sign of the current discrepancy is different from that of the discrepancies on the most recent previous trials on which the stimulus occurred.

Given the decreased weighting of older trials, a first order approximation to this equation would be merely changing salience of a stimulus present on a trial proportional to the product of the current discrepancy and the discrepancy on the most recent trial in which the stimulus previously occurred. Let us use this first order approximation to briefly consider the consequences of this theory for some simple cases of classical conditioning. For each case we assume the stimulus A occurs alone on every trial, and plot λ , the UCS intensity, and $\bar{V}=V_A$, the net associative

strength, versus time. For example, consider the case shown in Figure 1. Here the UCS occurs on every trial with equal intensity, and the initial salience is fairly low, so associative strength increases gradually. The discrepancy on a trial is immediately apparent in these graphs as the distance between the plots of λ and \bar{V} on that trial. That the discrepancy changes sign between two trials is indicated by the graphs of these two variables crossing each other, changing their relative positions, between those two trials. According to a first approximation to the theory presented here, this will result in a decrease in salience. In Figure 1, however, we see the opposite situation. The discrepancy remains large and of the same sign for several of the early trials. The first approximation to the theory presented here predicts that this will significantly increase α and thus will cause a positive acceleration in the early part of the learning curve. Later, as \bar{V} approaches λ , the current discrepancy and thus the increase in salience becomes much smaller, and the usual negative acceleration of the Rescorla-Wagner rule will predominate, producing an S-shaped acquisition curve overall.

The theory thus predicts that conditioning of a CS to a UCS will increase the salience of the CS, a result consistent with experimental observation. Note that although Mackintosh's theory (Mackintosh, 1975) also predicts an increase in salience in this case, it is for a very different reason. In that theory,

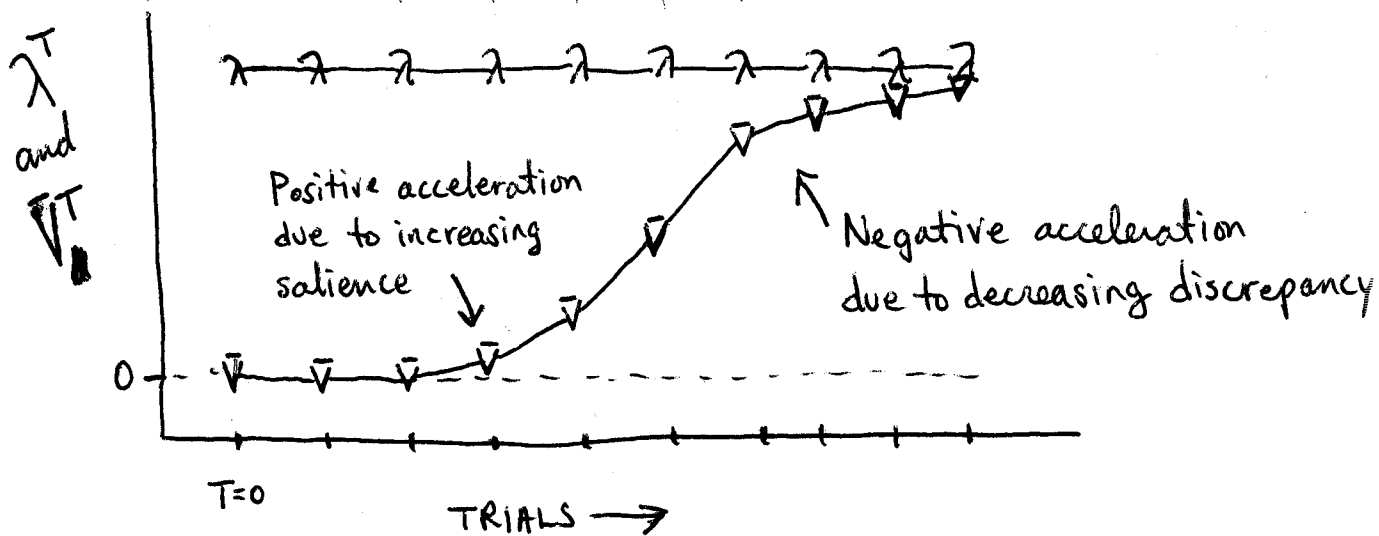


Figure 1. Predicted Time courses of net associative strength \bar{V} and reinforcement level λ as a function of trials for simple acquisition under a continuous reinforcement classical conditioning schedule. The greatest increase in salience is in the early trials, where the discrepancy is very large and consistently of the same sign. An early increase in salience can be ~~partly~~ ~~part~~ at least partly responsible for the positively accelerated initial portion of most acquisition curves.

salience would increase here due to the CS gaining a higher associative strength than any other cue (and thus having the smallest CS-specific discrepancy). In the theory presented here, on the other hand, salience is increased primarily before the CS gains strength, while the discrepancies are still large.

The theory also differs in its reason for making this prediction from the theories in which the salience of a CS increases by its being paired with reinforcement (e.g., Kirk, 1974; Frey and Sears, 1978). The theory presented here needs only a large discrepancy of unchanging sign. Thus, according to this theory, salience could be just as easily increased by pairing the CS with nonreinforcement when reinforcement was expected. For example, suppose stimuli B, C, D, ... were each strongly positively associated with a UCS, and they were one after the other presented in compound with neutral stimulus A in the absence of the UCS. The theory presented here would predict an increase in salience due to the consistent large negative discrepancy.

A second case, in which the initial salience is very high, perhaps $\alpha=1.5$, is shown in Figure 2. Even though reinforcement level is kept constant, the correction to the prediction \bar{V} will overshoot the actual reinforcement λ on each trial (see Figure 2). If α is not too high ($\alpha < 2.0$) \bar{V} will still converge to λ , as shown in the figure. So far, this is all a consequence of the

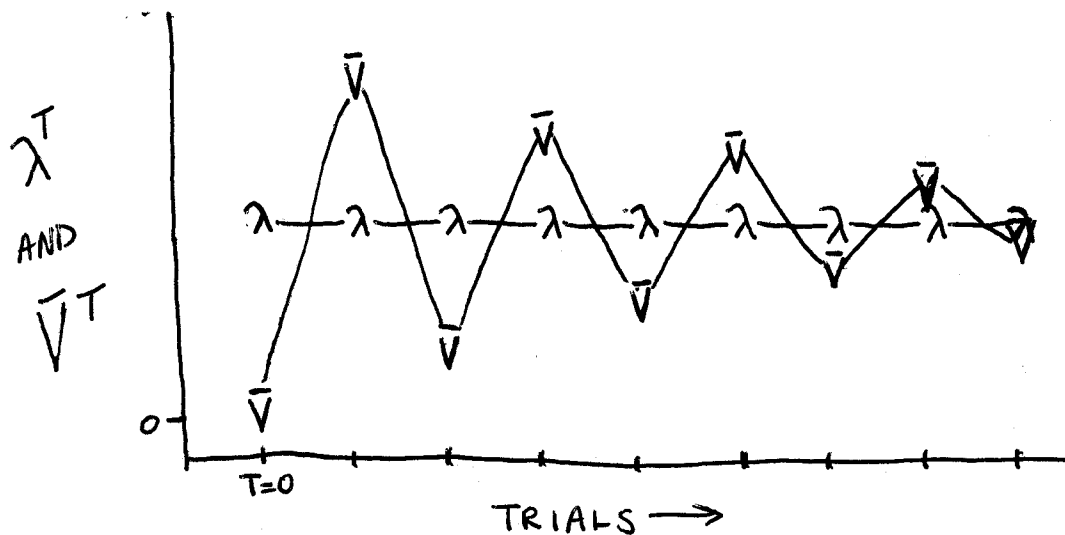


Figure 2. Net associative strength \bar{V} and reinforcement level λ as a function of trials under a continuous reinforcement schedule with very high initial salience.

Note that the discrepancy $D^T = \lambda^T - \bar{V}$ changes sign every trial. The form of this graph is purely a prediction of the Rescorla-Wagner theory alone. The salience change theory presented in this paper predicts ~~that~~ a large decrease in salience in this case.

Rescorla-Wagner rule alone. By looking at Figure 2, we can clearly see that the salience α is too large, that the consistent overshoot could be cured by reducing α . This is also the consequence of the present theory. Under these conditions the discrepancy actually changes sign every trial (as indicated by the crossing of the two graphs between every trial in Figure 2), and this is precisely the condition for salience to decrease.

By now the main idea behind the theory should be clear: If the discrepancies are in the same direction, as they are in Figure 1, then the associative strength is being repeatedly changed in the same direction, and performance will be improved if it is changed by larger amounts each time. On the other hand, if the discrepancies are repeatedly in opposite directions, then each change in the associative strength is having to be at least partially undone on the following trial, and this suggests performance could be improved by moving a little less far each trial. In order to capture this simple and important idea it is essential that the changes in salience depend on the relationship between discrepancies on different trials. No single trial contains the information necessary to make this sort of decision.

The two examples should also have provided some feeling for why we can be sure the salience value α will remain stable and bounded. For any reinforcement schedule there will be an optimal value for the salience that will minimize the mean square error

between \bar{V} and λ . Saliency following Equation 3 will always converge to that optimal value (for sufficiently small θ_A), whatever its initial value. For the above cases the optimal value was 1, since changing \bar{V} by exactly the amount of the discrepancy would have eliminated the error in one trial. This is not true when there is only a statistical relationship between CS and UCS, such as when a partial reinforcement schedule is used, and in this case the optimal value for saliency may lie between 0 and 1. Figure 3 shows the prediction of the theory for a CS on a 1/2 partial reinforcement schedule. Once the associative strength has reached its optimal level 1/2 of the way from $\lambda=0$ to $\lambda=1$, then each movement up or down following λ 's current value is actually counterproductive, and increases the likelihood of the discrepancy's changing sign. One can prove that under these circumstances the theory predicts saliency will become smaller and smaller, eventually asymptoting at zero. The fact that extinction is much slower in instrumental conditioning when a partial reinforcement schedule is used supports this prediction. As a symmetric consequence, the theory also predicts a retarded acquisition of a higher level of associative strength after a partial reinforcement schedule relative to a comparable group on a continuous schedule. To my knowledge, this has not yet been experimentally tested either in instrumental or classical conditioning.

The prediction of the theory of a decrease in saliency with

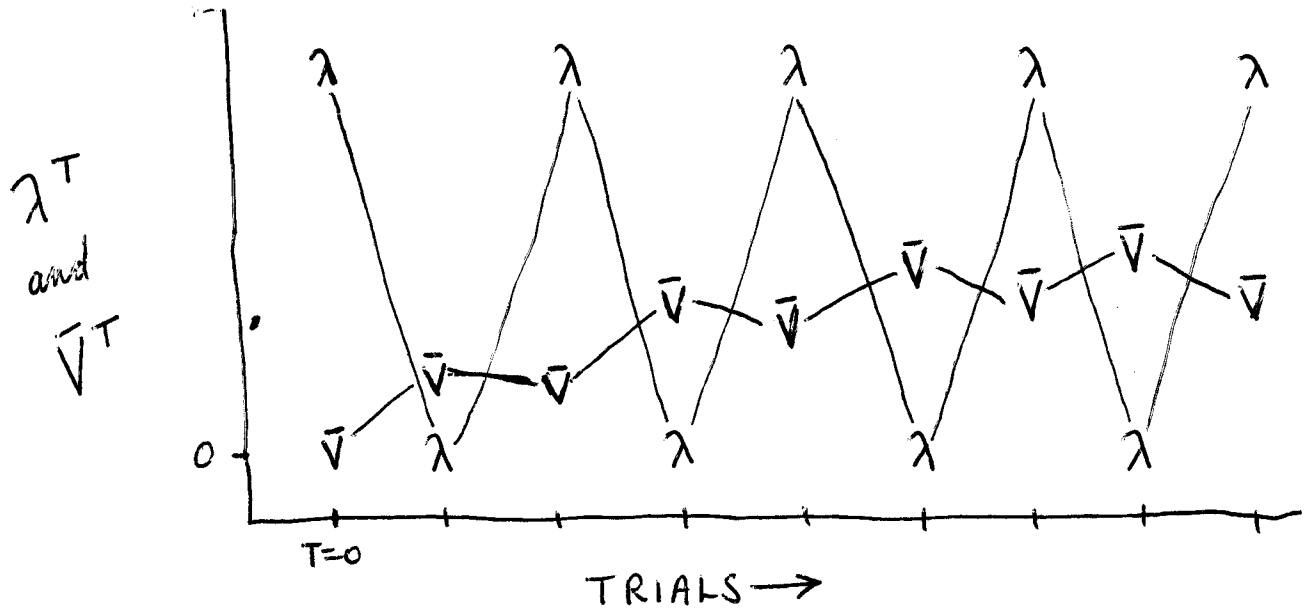


Figure 3. Prediction of Rescorla-Wagner rule for net associative strength \bar{V} and reinforcement level λ as a function of trials on a $\frac{1}{2}$ partial reinforcement schedule. The unpredictability of the reinforcement level tends to ^{cause} oscillation in the sign of the discrepancy once \bar{V} has reached its optimal level ~~the~~ $\frac{1}{2}$ of the way between the λ level on reinforced trials and the λ level on unreinforced trials. According to the theory of salience change presented in this paper, this will tend to cause a decrease in ~~the~~ salience towards zero.

a statistical relationship between CS and UCS still holds even if that relationship is one of statistical uncorrelation. In other words, salience should decrease if the CS occurs sometimes paired, sometimes unpaired with the UCS, even if the CS is uncorrelated with the UCS and provides no information as to whether or not it will occur. Because of this, the theory predicts that learned irrelevance or latent inhibition will occur if the pre-exposure period involves uncorrelated UCS presentations. If the UCS is not presented during the pre-exposure period, the theory has little basis for predicting a decline in salience.

Each of the parameters α , θ , and V have been used in this theory as being specific to a stimulus, but with no explicit mention of their response specificity. Each of these parameters should be thought of as being specific to a particular response, but also as being able to generalize to effect other responses. Thus, for example, conditioning a CS to a UCS will result in the greatest increase in the salience of the CS specific to that UCS, but will also increase the salience of the CS with respect to other UCS's. The theory presented here has little to contribute towards making the relationship between response specificity and response generalization more precise.

This theory of salience has been derived starting from the Rescorla-Wagner theory as its basis, but this is not strictly

necessary. A theory which captures much the same idea as the one presented here can be derived by assuming Mackintosh's rule for modifying associative strength (Mackintosh, 1975). The resultant salience change rule is just like Equation 3 except that $D_A = \lambda - V_A$ is used instead of Equation 2.

All current theories of salience change (Frey and Sears, 1978; Kirk, 1974; Mackintosh, 1975; Dickinson and Mackintosh, 1979; Moore and Stickney, 1980; Pearce and Hall, 1980) determine change in salience solely from the events occurring on a single trial. For example, in Mackintosh's theory (Mackintosh, 1975), stimuli present on a particular trial change their salience dependent on their relative discrepancies. The change in salience under this rule can be computed for a particular trial with no knowledge of what happened on any other trials. In the theory presented in this paper, on the other hand, salience change is completely dependent on the relationship between discrepancies on separate trials.

If the discrepancies on the trials in which a stimulus is present are always of the same sign, then the theory predicts that salience for that stimulus will be increased, so that the associative strength for that stimulus can be changed more rapidly to eliminate that consistent discrepancy. On the other hand, if the discrepancies on the trials on which a stimulus is present tend to oscillate in sign, then the theory predicts that

salience for that stimulus will be decreased, which in general will result in less oscillation and more accurate prediction.

This simple idea for improving prediction performance by changing salience requires comparing events on different trials in which the stimulus occurs, and thus has not been incorporated into any of the other current theories of salience. Of course, the question of whether or not animals actually use this idea remains for experimental test. The theory seems to be reasonably consistent with the experimental results already available, but a real test will come only with further experimentation. Several experimental predictions of the theory have been mentioned which differ from those of most other theories. In work in progress, the theory is being developed as a technique for improving the performance of prediction forming algorithms. The utility of the technique in artificial prediction systems suggests that natural ones may also make use of it.

BIBLIOGRAPHY

Dickinson, A., and Mackintosh, N. J. Reinforcer specificity in the enhancement of conditioning by posttrial surprise. Journal of Experimental Psychology: Animal Behavior Processes, 1979, 5, 162-177.

Frey, P. W., and Sears, R. J. Model of conditioning incorporating the Rescorla-Wagner associative axiom, a dynamic attention process, and a catastrophe rule. Psychological Review, 1978, 85, 321-340.

Kirk, R. J. Saliency modification of neocontinuity theory. Unpublished doctoral dissertation, Purdue University, 1974.

Mackintosh, N. J. A theory of attention: Variations in the associability of stimuli with reinforcement. Psychological Review, 1975, 82, 276-298.

Moore, J. W., and Stickney, K. J. Formation of attentional-associative networks in real time: Role of the hippocampus and implications for conditioning. Physiological Psychology, 1980, 8, 207-217.

Pearce, J. M., and Hall, G. A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of

unconditioned stimuli. Psychological Review, 1980, 87,
532-552.

Rescorla, R. A., and Wagner, A. R. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black and W. F. Prokasy (Eds.), Classical Conditioning II. New York: Appleton-Century-Crofts, 1972.

APPENDIX : Derivation of the Saliency Change Rule

Let us number all trials $t=1, 2, 3, \dots$, and indicate the trial number of each variable in ~~square~~ parentheses. $\bar{V}(t)$, for example, would indicate the net associative strength on trial t . Further, let us number all possible conditioned stimuli $i=1, 2, \dots, n$, and refer to stimulus number of a variable by subscripting. $\alpha_i(t)$, for example, would refer to the saliency of stimulus i at ~~time~~ trial t . ~~We now rewrite the Rescorla-Wagner equal rule in~~ ~~the following~~ Further, we introduce the variables $X_i(t)$, $i=1, \dots, n$, where $X_i(t)=0$ indicates stimulus i was not present on trial t and $X_i(t)=1$ indicates stimulus i was present on trial t . We can rewrite the Rescorla-Wagner rule in this notation as

$$V_i(t+1) = V_i(t) + \alpha_i(t) [\lambda(t) - \bar{V}(t)] X_i(t) \quad (1)$$

where

$$\bar{V}(t) = \sum_{i=1}^n V_i(t) X_i(t)$$

We assume the system is trying to minimize the square of the difference between predicted \bar{V} and actual reinforcement λ :

$$[\lambda(t) - \bar{V}(t)]^2 = D(t)^2$$

The second major assumption is that this minimization is done by each trial modifying the salience parameters $\alpha_i(t)$, $i=1, \dots, n$, ~~according~~ ^{according} ~~proportional~~ to how their change could have improved performance on that trial. To do this each salience parameter is modified proportional to the gradient of the error $D(t)^2$ with respect to that salience parameter:

$$\alpha_i(t) = \alpha_i(t-1) - c_i \cdot \nabla_{\alpha_i} D(t)^2$$

where $\nabla_{\alpha_i} D(t)^2$ indicates the gradient of $D(t)^2$ with respect to a change in ~~the~~ the value of α_i at all times. c_i is a small positive constant and the gradient is subtracted from the current ~~value~~ salience because we want to move down the gradient towards the minimum error. Now we merely solve for $-c_i \nabla_{\alpha_i} D(t)^2$:

$$\begin{aligned} -c_i \nabla_{\alpha_i} D(t)^2 &= \cancel{\text{messy}} - 2c_i D(t) \cancel{\text{messy}} \nabla_{\alpha_i} D(t) \\ & \quad \text{(by the chain rule)} \\ &= \theta_i D(t) \nabla_{\alpha_i} \bar{V}(t) \end{aligned} \quad (2)$$

where $\theta_i = 2c_i$ and the gradient is treated exactly like a partial derivative.

$$\nabla_{\alpha_i} \bar{V}(t) = \nabla_{\alpha_i} \sum_{j=1}^n V_j(t) X_j(t) = \sum_{j=1}^n X_j(t) \nabla_{\alpha_i} V_j(t)$$

which, by assuming $\nabla_{\alpha_i} V_j(t) = 0$ for $i \neq j$, we can approximate by

$$= X_i(t) \cdot \nabla_{\alpha_i} V_i(t) \quad (3)$$

$$\nabla_{\alpha_i} V_i(t) = \nabla_{\alpha_i} \left\{ V_i(t-1) + \alpha_i(t-1) [\lambda(t-1) - \bar{V}(t-1)] X_i(t-1) \right\}$$

(by Equation 1)

$$= \nabla_{\alpha_i} V_i(t-1) + [\lambda(t-1) - \bar{V}(t-1)] X_i(t-1) - \alpha_i(t-1) X_i(t-1) \nabla_{\alpha_i} \bar{V}(t-1)$$

(by the derivative of a product rule)

$$= \nabla_{\alpha_i} V_i(t-1) + X_i(t-1) \left\{ \lambda(t-1) - \bar{V}(t-1) - \alpha_i(t-1) X_i(t-1) \nabla_{\alpha_i} V_i(t-1) \right\}$$

(by Equation 3)

or,

$$\nabla_{\alpha_i} V_i(t+1) = \begin{cases} \nabla_{\alpha_i} V_i(t) & \text{if } X_i(t) = 0 \\ (1 - \alpha_i(t)) \nabla_{\alpha_i} V_i(t) + D(t) & \text{if } X_i(t) = 1 \end{cases}$$

Since $\nabla_{\alpha_i} V_i(t)$ clearly only changes on trials on which stimulus i occurs, we can easily rewrite this equation ~~to get a form~~ ~~indicating what~~ in terms only of the trials in which this stimulus occurs:

$$\nabla_{\alpha_i} V_i^{T+1} = D^T + (1 - \alpha_i^T) \nabla_{\alpha_i} V_i^T \quad (4)$$

where the superscript ~~T~~ indicates the trial number within the sequence of trials in which ~~the~~ stimulus i occurred. Thus trial T might be the current trial if ~~the~~ stimulus i occurred in it, ~~trial~~ ^{trial} $T-1$ the most recent previous trial in which stimulus i occurred, and trial $T+1$ the next trial in which stimulus i occurs. We can also combine Equations 2 and 3 in this notation to yield

$$\alpha_i^T = \alpha_i^{T-1} + \theta_i D^T \nabla_{\alpha_i} V_i^T \quad (5)$$

These two equations (4 and 5) are a recursive form for the salience change rule presented in the paper, as its Equation 3.