Simulation of habituation to simple and multiple stimuli

Emilio del Rosal, Lola Alonso, Rafael Moreno*, Manuel Vázquez and José Santacreu

Autónoma University of Madrid and * University of Seville

Address for correspondence:

José Santacreu

Department of Health Psychology, Autónoma University of Madrid.

Ciudad Universitaria de Cantoblanco. 28049 Madrid. Spain

E-mail: jose.santacreu@uam.es
Simulation of Habituation to simple and multiple stimuli

Abstract

Within the psychological literature there are a number of models that reproduce the defining properties of habituation to a single stimulus. However, most of them do not reproduce the phenomenon of dishabituation shown in empirical studies, consisting in the recovery of a stimulus previously habituated upon the appearance of a novel stimulus. The present work offers a model of habituation which, in addition to reproducing the basic properties of habituation to a stimulus, also does so when more than one stimulus is presented, and thus includes the dishabituation phenomenon. This model consists of two functions, one called “activation” and the other “availability”, and is tested by means of simulation of the responses in the context of different stimulus patterns. The results of the simulation show a good qualitative fit to the empirical results on the phenomena of habituation, including dishabituation. In addition, the model is suitable for inclusion in associative models that reproduce classical conditioning, which will make it possible in the future to incorporate into these the influence that the habituation of each stimulus may have on its association with other stimuli.

Keywords: Dishabituation; Habituation; Models; Multiple stimuli; Simulation
The phenomenon of habituation is defined, in contrast to other response decrements, through diverse properties, among which are the following (Groves & Thompson, 1970; Thompson & Spencer, 1966):

1. Exponential decrease of the strength of the response with repeated exposure to a stimulus.
2. More rapid habituation with stimuli of lower intensity.
3. More rapid habituation with shorter intervals between stimuli (ISI).
4. Spontaneous recovery of the response after a period without exposure to the stimulus.
5. Recovery of the response or dishabituation of the stimuli previously habituated upon the introduction of a new stimulus.

With these properties, habituation has a significant function consisting in “filtering” stimuli that are irrelevant for subjects, so that they do not have to respond to them (Hall, 1991, for a fuller account).

Various models have been proposed for reproducing the properties of habituation. After a first computational model by Stanley (1976) on short-term habituation, that of Innis and Staddon (1989) reproduced the characteristic negative exponential curve and dependence of the speed of habituation on ISI. The proposals of Wang and Hsu (1990), Wang and Arbib (1992) and Wang (1994) focus on combining short- and long-term habituation in a single model. To this end they pair two differential equations. One of these refers to the activation of the stimulus in each presentation, and the other describes changes to speed of recovery with an inverse S-shaped curve, whose high initial values and low final values correspond to the two types of habituation. This permits the researcher to estimate the response to the presentation of a stimulus and simulate the typical habituation curve. With a different approach, two later studies (Staddon, 2002; Staddon & Higa, 1996) introduce different linked processing units, which, according to their peripheral or central position in the chain, intervene in the short or long scales of time between stimuli. This has made it possible also to reproduce the sensitivity of spontaneous recovery to the inter-stimulus interval (ISI), an aspect no other model has managed to achieve –including the one we presented in a previous work of our group (Alonso, Moreno, Vázquez and Santacreu, 2005).

In sum, the more advanced models proposed by the groups of Staddon and of Wang are capable of reproducing a large part of the phenomena characteristic of habituation. However, none of those models take into account exposure to more than one stimulus. Therefore, they do not consider reproducing...
Simulation of Habituation

experimental data such as those presented in Epstein, Rodefer, Wisniewski and Caggiula (1992), Post and von der Emde (1999) or Wiel and Weeks (1996), which include at least a second stimulus, and show the phenomenon of dishabituation. This phenomenon consists in the recovery of a habituated stimulus, $S_1$, upon presentation of a new stimulus, $S_2$; this habituation then follows its course, rapidly attaining levels similar to those prior to presentation of the second stimulus.

It would be desirable, therefore, to have access to a model of habituation which reproduced the basic properties of habituation to more than one stimulus, included dishabituation to a stimulus through the introduction of a new one. We have not found such an integrative model in the literature, nor even the topic “habituation to multiple stimuli”. We can point to some exceptions, at least partial: the model of Schmajuk, Lam and Gray (1996), which combines aspects of classical conditioning and habituation, or the SOP model initially proposed by Wagner (1981) and developed by Brandon, Vogel and Wagner (2003) and Vogel, Brandon and Wagner (2003). These models compute various stimuli but place the emphasis on the study of conditioning, and do not cover reproduction of the basic properties of habituation, such as variations according to ISI or dishabituation. On the other hand, the model of Lara (1983) considers habituation to two stimuli, but deals with the matter from the perspective of neuronal networks with different physiological layers, an approach that differs from those of the models mentioned above as references for the one we shall present. Moreover, Lara’s model does not function in real time, and nor, therefore, does it adjust to the stimulus-related changes that occur as a result of previous events; it requires a priori definition of the values of diverse initial variables, such as interval between stimuli and between different series of them, number of times each stimulus is presented, its intensity, and other parameters.

The present work proposes a model of habituation capable of describing the principal characteristics of the phenomenon both for one stimulus and more than one, including, therefore, dishabituation to one of them upon exposure to a new stimulus. This model basically rests on the idea of two paired functions proposed by the models of Wang and colleagues. Together with a function for decay of the response, the model presented adopts an idea analogous to that of Wang (1994) in relation to recovery, also proposed by Staddon and Higa (1996). It opts for iterated rather than differential functions, and substitutes the inverse sigmoidal function constructed by Wang’s group of recovery rate with an increasing sigmoidal function of availability. Furthermore, the model incorporates two elements that are fundamental to its objectives. The first of these is the notion of “stimulus trace” proposed for associative models by Sutton and Barto (1981); understood as residual activation that persists for a period after appearance of the stimulus, this notion is useful for considering the influence of the previous state of the stimulus on habituation. The second element is a new term with two objectives: to manage multiple stimuli and to make it possible for the values of each
Simulation of Habituation

one to form part of the equations of the rest of the stimuli, thus influencing their availability.

Method

Components of the Model

The model estimates the value of the responses to each stimulus, and consists of two inter-related functions for each stimulus $j$, activation, $A_j^t$ and availability, $D_j^t$. Activation refers to the values of the organism’s response to the presentation of a stimulus of a given intensity, while availability regulates activation, especially upon repeated presentation of the stimulus. Both are iterated functions, which means that, based on the specification of an initial condition $f(0)$, each one of their values is always a function of the previous one; in other words, $f(n+1) = g(f(n))$. In our model, as in that of Wang and colleagues, each iteration or step for the calculation of functions corresponds to a time-cycle $t$ that represents one second, incorporated into the expression of the functions as $A_j^t$ and $D_j^t$. Let us consider each one in detail and in a progressive fashion.

The function $A_j^t$

The values of $A_j^t$ depend on the intensity of the stimulus $S_j$, with $S_j = 0$ when the stimulus is absent and $1 > S_j > 0$ when it is present, and also on the values of $D_j^{t-1}$. As shown by the following expression

$$A_j^t = S_j^t \cdot D_j^{t-1} \quad (1)$$

the increase in activation produced by a stimulus $j$ is simply its intensity weighted by the existing availability. Thus, at the moment of presentation of a stimulus, the value of activation will be greater the higher the value of $D_j^t$. Bearing in mind the possible values of $S_j$ indicated above, and that the function $D_j^t$ must lie between 0 and 1, the values of $A_j^t$ are also between these two values, except for certain cases in which sensitization occurs.

In order for the activation to decay gradually with passage of time without further presentation of the
stimulus, we include a new term, $\lambda \cdot A_j^t$, where $0 < \lambda < 1$, giving the following function:

$$A_j^t = \lambda \cdot A_j^{t-1} + S_j^t \cdot D_j^{t-1}$$  \hspace{1cm} (2)

In this way, at the time of exposure to the stimulus, activation will increase sharply due to the term $S_j^t \cdot D_j^{t-1}$. When the stimulus $j$ is absent, and therefore $S_j^t = 0$, the second term of the equation is 0, and $A_j^t$ will necessarily be less than $A_j^{t-1}$ given its multiplication by $\lambda$. Thus, if the stimulus is not presented anew, activation will decay exponentially, as proposed. Following Sutton and Barto (1981), we use the term "stimulus trace" to refer to this residual activation that persists after presentation of the stimulus, obtained by the term $\lambda \cdot A_j^{t-1}$.

As regards the response $R_j^t$ this is formally defined as:

$$R_j^t = \begin{cases} A_j^t & \text{if } S_j^t \neq 0 \\ 0 & \text{if } S_j^t = 0 \end{cases}$$  \hspace{1cm} (3)

which indicates that the response will appear after each presentation of the stimulus with a value equal to $A_j^t$.

**The function $D_j^t$**

The function $D_j^t$ included in the Activation function (Equation 2) represents the potential activation of the stimulus at the time of its presentation, has a positive sigmoidal form between 0 and 1 and is useful for making it possible for activation to decrease the more a stimulus is repeated. A function with sigmoidal form and expressed in an iterated manner would be as follows:

$$P_j^t = P_j^{t-1} + r \cdot P_j^{t-1} \cdot (1 - P_j^{t-1})$$  \hspace{1cm} (4)

where $r$ determines the slope of the curve, i.e., the speed with which it increases. Our function $D_j^t$ is the one
Simulation of Habituation shown, with \( D'_j = P'_j \), with slight modifications and the consideration of the two following calculation conditions.

\[
D'_j = \begin{cases} 
  D_{j-1}' + \left( \kappa \cdot \frac{A_{j-1}' \cdot D_{j-1}'}{S'_j} + \beta \cdot \sum S'_n \cdot A'_n + \rho \right) \left( 1 - D_{j-1}' \right) & \text{if } S'_j = 0 \\
  \xi \cdot \left( \frac{A_{j-1}' \cdot D_{j-1}'}{S'_j} + \varepsilon \right) + \beta \cdot \sum S'_n \cdot A'_n & \text{if } S'_j \neq 0 
\end{cases}
\]  

(5)

We shall begin with the expression in the case that \( S'_j = 0 \). \( D'_j \) is a sigmoidal function in which the term \( r \cdot P_{j-1} \) is substituted by another with more elements to be able to determine the speed of increase or recovery of availability \( D'_j \).

This added term

\[
\kappa \cdot \frac{A_{j-1}' \cdot D_{j-1}'}{S'_j} + \beta \cdot \sum S'_n \cdot A'_n + \rho
\]  

(6)

contains the following components. The parameter \( \kappa \) regulates the speed of recovery of availability, just like \( r \) in the considered expression of the sigmoidal function. In the term \( \frac{A_{j-1}' \cdot D_{j-1}'}{S'_j} \), the numerator includes the factor \( A_{j-1}' \) for which availability increases more rapidly in the initial phases of habituation, when the activation still has high values; the higher the \( A_{j-1}' \), the more rapidly the availability will increase. The interaction between the two terms of the numerator means that stabilization in the asymptote will occur more slowly, as occurs with empirical data. The denominator \( S'_j \) serves to avoid stimuli of extreme intensities (high or low) producing inappropriately slow or rapid habituations. In order to prevent the denominator of the mentioned expression being null, the following calculation condition was previously defined:

\[
S''_j = \begin{cases} 
  S'_j & \text{if } S'_j \neq 0 \\
  S''_{j-1} & \text{if } S'_j = 0 
\end{cases}
\]

In the following component of the enlarged term, the result of the expression \( \beta \cdot \sum S'_n \cdot A'_n \) reflects the
activations of the set of stimuli that can be present, reflecting the “trace” of each one of them. Thus, when these display non-negligible activations, increase of $D'_j$ will rise, and therefore also the values of $A'_j$ in subsequent presentations of the stimulus –and this is exactly in what dishabituation of stimuli upon presentation of another stimulus consists. This value is regulated, in turn, by the parameter $\beta$. Finally, the parameter $\rho$ adds its value to the general expression and permits $D'_j$ to continue increasing even when the values of $A'_{j-1}$ approach 0 and no other stimuli are present.

When the stimulus is present, $S'_j \neq 0$, availability falls sharply by virtue of the values close to 0 of the parameter $\xi$. The parameter $\varepsilon$ indicates the point of equilibrium of a stimulus when it is presented in a continuous fashion, permitting the non-null value of the asymptote of the habituation.

In sum, when the stimulus is absent and $S'_j = 0$, the availability increases in sigmoidal fashion with an asymptote with a value of 1; on the other hand, when the stimulus is present and $S'_j \neq 0$, the availability must undergo a sharp fall and require some time before being able to recover fully.

As regards the parameters used, $\lambda$, $\beta$ and $\varepsilon$ are freely defined, while the rest, $\kappa$, $\rho$ and $\xi$, depend on $\lambda$ and in some cases also $\varepsilon$, in accordance with expressions estimated and calibrated previously (Alonso, Vázquez & Santacreu, 2003), which can be consulted in the Appendix.

**The dynamics of the model. The interaction between $A_{j-1}'$ and $D_{j-1}'$**

Figure 1 shows the behaviour of the two functions in key circumstances for different properties of habituation, namely, repeated presentation of a stimulus, non-presentation of a stimulus for a period and the difference for different intervals between stimuli –the key circumstances for properties 1, 4 and 3, respectively, of those listed at the beginning of this article. For this we programmed, respectively, 5 presentations of a stimulus with an interval of 10 seconds, a rest period of 70 seconds and 3 further presentations, this time with an interval of 45 seconds. For the correct functioning of the model, the initial values of $A'_j$ and of $D'_j$ are 0 and 1, respectively; activation of the stimulus should start out from 0, while its availability must be maximal, as the stimulus has not been presented.

**INSERT FIGURE 1**
Simulation of Habituation

As it can be seen, with each presentation of the stimulus the relative peaks of activation have lower and lower values, as is characteristic in habituation experiments. Availability shows a sharp fall with each presentation of the stimulus, subsequently recovering in sigmoidal fashion. The key issue of $D_j$, namely, that its recovery progressively declines, and thus also its contribution to the peaks of $A_j$, is due (see equation 2) to the fact that the slope of $D_j$ is directly proportional to $A_j$, whose values are themselves decreasing in the case considered here. It is, therefore, the connection between the two functions that produces the behaviour characteristic of habituation. Likewise, the passage of time without further presentations of the stimulus permits the recovery of $D_j$, so that the value of the activation in a subsequent new presentation rises slightly with respect to that of the final presentations prior to the rest period. Thus, the recovery of the response characteristic of habituation is reproduced. Moreover, the higher ISI of the final trials produces gentler curves than at the beginning, when the ISIs were lower.

Figure 2 shows the curves of the two functions $A_j$ and $D_j$, generated now by a single presentation of two stimuli, $S_1$ and $S_2$, of different intensities, 0.50 and 0.90, respectively –property 2. The lower-intensity stimulus provokes lower activation, achieved because the intensity $S_j$ weights the effect of $D_j^{-1}$ on $A_j$ (see equation 1). Also of note is the recovery of the availability of the first stimulus $S_1$ upon presentation of the second one $S_2$, indicated above as property 5, due to term $\beta \cdot \sum S_n \cdot A_n^j$, which includes the consideration of more than one stimulus, and therefore the traces of each one of them.

**INSERT FIGURE 2**

**Procedure**

The model described was tested by applying it to three different stimulus patterns in which we could appreciate the properties of habituation it was intended to reproduce. The patterns are similar to those described in experiments such as those of Broster and Rankin (1994), Epstein, Rodefer, Wisniewski and Caggiula (1992), Post and von der Emde (1999), Rankin and Broster, (1992), Rankin and Wicks (2000), or Wiel and Weeks (1996). Given that the quantitative values of the responses to habituation vary from one study to another as a function of the parameters and type of subjects employed, we assess the simulation in a qualitative manner, reproducing the characteristic form of the habituation and dishabituation curves. The stimulus patterns considered were as follows:
Simulation of Habituation

Pattern A: In a first phase, a stimulus of intensity 0.95 is presented on 30 occasions with ISI of 15 seconds, leaving 350 seconds of rest before a final presentation. The aim of this is to simulate the exponential diminution of the response with repeated presentation of a stimulus, and also the recovery upon a new presentation of this S at some time after the previous presentation –properties 1 and 4 of those listed at the beginning of the paper.

Pattern B: A single stimulus of intensity 0.99 is presented three times on 30 occasions, with ISI of 10, 30 and 60 seconds, to reveal whether more rapid habituation is produced with shorter ISIs (property 3).

Pattern C: In order to test the behaviour of the model with multiple stimuli, one of them, $S_1$, of intensity 0.95, is presented several times with an ISI of 30 seconds. At the mid-point between presentations 20 and 21, a second stimulus, $S_2$, of similar intensity, is presented just once to reveal whether dishabituation of stimulus $S_1$ is produced (property 5). All of this occurs in the context formed by a stimulus $S_3$ also presented from the beginning, and repeatedly, with constant intensity 0.1, and for which, in turn, we should be able to appreciate the dishabituation produced by the presentation of $S_2$.

Finally, it should be pointed out that in order to favour the desired behaviour of the model, it was decided to use the values indicated by the free parameters, $\lambda = 0.95$, $\beta = 0.5$ and $\epsilon = 0.01$, the following ones thus being obtained for the rest, according to the expressions shown in the Appendix:

$$\xi = 0.025, \quad \kappa = 0.3171 \quad \text{and} \quad \rho = 0.03804.$$

Results

Figure 3 shows the results of applying the model to pattern A, as described above. The data correspond to the maxima of the activation function. The reader can see first of all the descending curve characteristic of habituation upon repeated presentation of a stimulus S, and also the recovery from that habituation when some time later the stimulus is presented anew. On this latter occasion, the level of response does not attain the initial levels.

Figure 4 shows the behaviour of the model in response to pattern B as described above. The diminution of response characteristic of habituation occurs more rapidly when ISIs are shorter. With these shorter ISIs, the asymptotic values of the activation are lower, and the curves of the habituation slope are
more pronounced, reflecting more rapid decrease. It can also be seen that with certain ISIs, there is
sensitization of the stimulus in the presentations immediately following the initial one, as observed in the
literature (e.g., in Rankin & Broster, 1992). This effect is achieved by the model through different possibilities.
With high values of $\lambda$, there will probably be sensitization for high ISIs, since the stimulus trace $\lambda \cdot A_{j}^{t-1}$ will
have diminished little; it will also occur on increase of the availability function when the values of $\kappa$ are high.

INSERT FIGURE 4

Figure 5 shows the behaviour of the model in response to pattern C, which includes three stimuli. In
order to study this new situation in more detail, in Figure 5 we show all the values of the activation function,
rather than (as was the case in Figures 3 and 4) just its maxima. It can be appreciated that the stimulus $S_1$
initially habituated is dishabituated upon presentation of the second stimulus $S_2$. This dishabituation
response is smaller than the initial one, as is usually reported in the literature. A similar effect, and
proportional to its intensity, is observed in $S_3$, the stimulus that represents the context of the two stimuli
considered as central.

INSERT FIGURE 5

Conclusions

The results obtained appear to indicate that the model constructed satisfactorily reproduces the key
properties of habituation, whether just one stimulus or more are presented. The inclusion of more than one
stimulus has made it possible to reproduce the dishabituation to one of them upon presentation of another.
This suggests the relevance of the involvement of the different components of the model, endorsing its
efficacy. The objective of the present work is therefore fulfilled, insofar as it deals with the lack of models for
habituation of the response with more than one stimulus.

In addition to the benefit for habituation, the present study establishes a new line of connection
between two fields previously insufficiently linked: habituation with multiple stimuli and processes involving
association between two or more stimuli. As we understand it, a model of habituation that can compute
several stimuli is the condition for integrating habituation and associative learning. In our view, some aspects
of such learning can only be dealt with if we include the consideration of the process of habituation of the
stimuli associated, bearing in mind that associability of stimuli also depends on their activation, through the stimulus trace, at each point in time.

With this purpose, Alonso, Moreno, Vázquez, del Rosal and Santacreu (2005) designed the model proposed here with a series of functions of association between stimuli. This permitted consideration in a simulation of classical conditioning of how spontaneous recovery of the conditioned response was influenced by habituation to the participating stimuli. This study supports the idea that models of learning should not ignore less complex processes such as habituation and, moreover, shows the capacity of the presented model for incorporating functions that allow consideration of the degree of association between the multiple stimuli that may be present.

References


Figure caption

Figure 1: Behaviour of the functions $A_j$ and $D_j$ with repeated presentation of a stimulus, a rest period and further presentations with higher ISI.

Figure 2. Behaviour of the functions $A_j$ and $D_j$ upon presentation of two stimuli with different intensities, 0.50 and 0.90.

Figure 3. Function $R_j$, or maxima of $A_j$, at successive moments of presentation of a stimulus, with ISI of 15 seconds, and after a period without its presentation.

Figure 4. Habituation curves to stimuli of equal intensity and different ISIs, 10, 30 and 60 s.

Figure 5. Dishabituation of stimulus $S_1$ upon presentation of a second stimulus $S_2$, in a context defined by $S_3$.
Expressions of the parameters $\kappa$, $\rho$ and $\xi$ used, estimated and calibrated by Alonso, Vázquez and Santacreu (2003).

\[
\kappa_{\epsilon=0} = \left( \frac{1}{\lambda^2} + \lambda - 2 \right) \cdot 5.5
\]

\[
\frac{1}{\kappa_{\epsilon \neq 0}} = \frac{1}{\kappa_{\epsilon=0}} + \frac{1}{2.85 - 2.34\lambda} \cdot \epsilon^{\left( \frac{0.75+0.33\lambda}{1+2.67(\lambda-\lambda^2)} \right)}
\]

\[
\rho = 0.0118 \cdot (1-\lambda)^{0.36} \cdot 0.36^{(1-\lambda)} - 0.02 \cdot (1-\lambda) \cdot \epsilon
\]

\[
\xi = 0.5 \cdot (1-\lambda)
\]

\[
\kappa_{\epsilon=0} = \left( \frac{1}{\lambda^2} + \lambda - 2 \right) \cdot 5.5
\]

\[
\frac{1}{\kappa_{\epsilon \neq 0}} = \frac{1}{\kappa_{\epsilon=0}} + \frac{1}{2.85 - 2.34\lambda} \cdot \epsilon^{\left( \frac{0.75+0.33\lambda}{1+2.67(\lambda-\lambda^2)} \right)}
\]

\[
\rho = 0.0118 \cdot (1-\lambda)^{0.36} \cdot 0.36^{(1-\lambda)} - 0.02 \cdot (1-\lambda) \cdot \epsilon
\]

\[
\xi = 0.5 \cdot (1-\lambda)
\]