1 Habituation Techniques

Christopher T. Fennell

Abstract

This chapter presents the general aspects of the habituation technique. This technique has helped to address various language acquisition questions over the past half-century. While discussing implementations using different behavioural responses, the chapter focuses on the most common measure of habituation in language acquisition research: looking time (LT). Issues in implementing the method and potential problems are discussed. The simplicity of the habituation procedure in both its design and implementation, along with its long history in the field, makes this method one of the fundamental tools that psycholinguists can use to uncover nascent, emerging, and maturing language skills during infancy and early childhood.

Assumptions and Rationale

One of the biggest challenges of determining what an infant knows about language is actually tied to language itself. Unlike Piaget (1926), who famously asked older children to reflect on and discuss their understanding of the meaning of words, we have no such luxury of interviewing a 12-month-old regarding their word-referent links. Even for developmentally simpler skills, we cannot get a 6-month-old to give a simple yes or no answer to the question of whether they discriminate two language sounds. It is somewhat paradoxical that language itself is a barrier to understanding
language development in infants. The fact that infants have little or no lexical production requires researchers to often turn to tasks that require no language output from the child. Further, infants’ limited motor skills restrict the measures that can reveal underlying linguistic abilities. Tasks must take advantage of gross motor abilities, such as full head turns (fine pointing or manual selection are difficult); congenitally organized behaviors that infants have strong control over since birth, such as looking or sucking; or, basic psychophysiological responses, such as heart rate.

One of the most valid and reliable tools we have to examine the perceptual skills related to infant language is the habituation task. Habituation is a decrease in a response to a stimulus after repeated presentations. This produces what is termed the habituation curve, a monotonically decreasing behavior in response to a repeated target stimulus. It is a task with a very long history in our field, stretching back to the nineteenth century (for a review, see Thompson, 2009). Indeed, Thompson highlights that the concept is reflected in antiquity: in Aesop’s fables, a fox is quite frightened of a lion upon first meeting him, but becomes less alarmed upon each subsequent viewing. Perhaps Aesop’s example was prescient. Habituation tasks were primarily used for decades with animals (and continue to be used with these populations), with everything from amoebas to dogs showing habituation responses (Harris, 1943).

Considering the long history of the task and ubiquitous nature of the habituation response across other non-verbal beings (i.e., animals), it is unsurprising that the method was extended to infants in the early twentieth century (see Humphrey, 1933). However, simply habituating an infant to a stimulus is necessarily a bit limiting with respect to what one can say about learning. If, for example, an 8-month-old had a reduced behavioral response to the repeated presentation of a phoneme, one could argue that they have formed a memory of that particular sound. But it could also be that the infant is simply tiring. The key to demonstrating that the infant has formed a representation of or learned something about the presented stimulus is dishabituation—an increase in behavioral response to a novel stimulus.

Sokolov’s (1963) comparator model is the classic formulation of this approach. The infant (or adult) has an orienting response to a novel or unexpected non-threatening stimulus (e.g., becoming still, looking at the stimulus, reduced heart rate). As it repeats, the infant builds an internal representation of the stimulus. The increasing strength of the representation leads to a greater match between the internal percept and the repeating external stimulus. The initially large orienting response correspondingly reduces as the internal/external match increases. But, if the external stimulus does not match the established internal representation (i.e., a novel stimulus), the infant’s orienting response should reoccur.

Thus, habituation is one of the optimal tasks for testing pre-verbal infants as it does not rely on overt productions, but rather on implicit cognitive measures such as those mentioned earlier (e.g., looking time, sucking, heart rate, among others). Further, based on the comparator model, it allows researchers to determine the nature of infants’ percepts and concepts by testing differing levels of novelty from the habituated stimulus (e.g., changing a habituated word form by one phoneme, or multiple sound changes). If the infants’ behavioral responses increase to the novel stimulus, it can be concluded that they have the ability to differentiate the habituated and novel stimuli. In this regard, habituation is fundamentally a method to index discrimination ability.

Fantz’s (1964) article in *Science* on visual habituation in the human infant broadly introduced using this task with very young participants to psychological researchers. However, it is important to note that previous studies had already used habituation with
infants, including studies on auditory habituation. For example, Bartoshuk (1962) demonstrated that newborns habituate to tones and dishabituate to tones of a differing intensity, using heart rate as his dependent measure. Once it was determined that infants could habituate and dishabituate to auditory tones, it was a straight road for researchers to examine language sound (i.e., phoneme) discrimination using similar methods.

In one of the seminal works on infant language perception, Eimas, Siqueland, Jusczyk, and Vigorito (1971) used a habituation task with sucking as their measure to investigate 1- and 4-month-old infants’ discrimination of consonants, specifically a voicing contrast. Consonants produced in the same place and manner can differ in the timing of the vibration of the vocal folds. For example, /b/ and /p/ are both produced from the lips and are stops, but they differ in voicing. Vocal cord vibrations occurring approximately 25 ms after the air burst from the mouth (or later) sound like a /p/ to English speakers. Vibrations starting before that mark sound like a /b/. In Eimas et al., infants heard a repeated sound contingent on strong sucks. Once their sucking rate decreased by 20%, a novel stimulus was presented in two experimental conditions. In one condition, the novel sound was from the same phonological category (i.e., a new /b/ sound that differed by 20 ms in voicing from the original stimulus) and in the other condition the novel sound came from a different category (i.e., a 20 ms voicing change that crossed the boundary from /b/ to /p/). In the control condition, the same sound was played after the 20% reduction in behavior. Only infants in the differing category condition had a dishabituation response—increased sucking when the sound change occurred.

The above experiment highlights some important aspects of infant habituation. First, habituation allows the researcher to test categorical perception in that we can determine if an acoustically different stimulus will engender a continued habituated response or a dishabituation response. As Thompson and Spencer (1966) highlight in their classic list of the characteristics of habituation, “habituation of response to a given stimulus exhibits stimulus generalization to other stimuli” (p. 19). Thus, we can assume that the lack of dishabituation to an acoustically different stimulus means that the infant considered it to fall into the same category, or the distinction is too weak to detect. This second explanation is unlikely if you include a condition where a similar magnitude difference elicits a dishabituation response due its crossing of a category boundary, as in the Eimas et al. work.

Second, the use of criterion equated the processing of the stimuli across infants. Based on Sokolov’s (1963) theory, the reduction in the target behavior is commensurate with the increasing robustness of the infant’s memory trace for the stimulus. But, different infants would potentially, and probably, have differing timing with respect to the building of the stimulus memory trace due to individual differences in cognitive skills, particularly attention. By requiring them to reach the same relative decrease, the researcher can assume that they have reached similar processing levels for the stimulus in question.

**Apparatus**

Testing infants’ language skills via habituation typically involves relatively little technology in comparison to some other methods present in the literature, like event-related potentials (ERP), functional near infrared spectroscopy (fNIRS), or functional magnetic resonance imaging (fMRI). As such, the necessary apparatus can be implemented relatively quickly and inexpensively in the lab. Primarily, a
researcher requires devices to present stimuli in a controlled manner and to measure
the target behavior. As the measurements of the latter need to feed back to control
the former in habituation, we typically use the same device to do both: habituation
software on a lab computer. A widely used free-ware program is available for such
purposes, called Habit 2 (Oakes, Sperka, & Cantrell, L., 2015). The program will
control stimuli presentation, compute habituation criteria, and accumulate
behavioral data. Stimuli are usually played from digitized files on the computer and
are sent to the display and speaker in the testing room. The experimenter, who
should be blind to the audio stimuli being presented and to whether a trial was a
habituation or test trial, remotely monitors the infant’s behaviors via key presses.

As alluded to in the section above, many of the early studies in infant habituation
used sucking or heart rate as the dependent measure. Measuring infant sucking
strength and rates requires the experimenter to have a pressure transducer within a
pacifier, and the corresponding connected equipment to measure the output from the
transducer. Heart rate measures usually require three electrodes to be placed on the
infants’ chest and abdomen, with the electrodes again connected to equipment to
measure their output. While these measures are still used, the typical behavior being
measured in modern habituation research is looking time (LT) to a visual display,
even if one is testing habituation/dishabituation of auditory language stimuli. There is
a positive relationship between attention to an auditory stimulus and visual fixation
(Horowitz, 1975). Unlike the measures above, LT requires nothing to be in physical
contact with the infants, which is an advantage. The researcher—appropriately
blinded to the condition—simply needs to record, by pressing buttons on a keyboard
connected to the same software, where the infant is looking, typically through watch-
ing the infant via a closed-circuit video camera. The use of a video camera also allows
for a record that can confirm the real-time measurements when coded post-experiment.
The ease of measuring looking behavior has led to its wide application.

Nature of the Stimuli

Due to the nature of the task, many habituation studies investigating language
development have involved basic auditory stimuli, such as changes to the acoustic
form of simple syllables. For example, tracking infants’ phonetic and phonological
development has been the focus of many infant habituation studies, starting with
Eimas et al. (1971). Using an example from Polka and Werker (1994), infants in such
studies are habituated to one syllable (e.g., /dYt/) and then given a syllable involving a
single phoneme change for the novel stimulus at test (e.g., /dut/). If infants dishabitu-
ate to the novel syllable, they are able to distinguish the target phonemic contrast.
See Figure 1.1 for a visualization of such a study. Such studies contributed to the
finding that infants are initially universal listeners, able to distinguish sounds from
both their native and from non-native languages, but then become language-specific
listeners over the first year—failing to dishabituate to non-native contrasts.

Habituation studies examining language development are not limited to simple sylla-
bles, however. For example, Mehler, Jusczyk, Lambertz, Halsted, Bertoncini, and Amiel-
Tison (1988) used a habituation method where the target stimuli were narrative auditory
passages from rhythmically similar and dissimilar languages (recorded from fluent
bilinguals so that the voice did not differ). Using sucking rate as the dependent variable,
they showed that infants of 2 months can distinguish their native language from a non-native language based on their rhythmic class, but not two non-native languages.

One can even use visual stimuli to demonstrate language discrimination. In a novel twist, Weikum, Vouloumanos, Navarra, Soto-Faraco, and Sebastián-Gallés (2007) presented infants with silent video clips of fluent French-English bilinguals reciting passages in each language. Infants of 4 and 6, but not 8, months dishabituated to French clips after being habituated to English ones, and vice versa. This shows that infants have an early ability to visually discriminate non-native from native languages before perceptually narrowing to their native language in the visual domain. Interestingly, French-English bilingual infants, for whom both languages were native, were able to discriminate the languages at the older age.

Finally, rather than focusing on audio or visual stimuli, some habituation experiments explore the connection between the two by pairing objects and word forms during habituation, and then test infants on novel word-object associations (see Figure 1.1). As such, these studies contribute to a major area of early language development: early word learning. For example, a simple way to invoke word learning is to replace a visual pattern typically used in discrimination studies with an object that affords naming. But infants may succeed by ignoring the object and simply focus on the change in label during the novel trial in the test phase. Werker, Cohen, Lloyd, Casasola, and Stager (1998) corrected for this by creating an audio-visual variation of the habituation procedure called the Switch task (see Figure 1.1). Infants are habituated to two word-object associations (e.g., Object A – Word A; Object B – Word B) and then tested on two trials: a Same trial comprising of one of the habituated pairings (e.g., Object A – Word A) and a Switch trial where an incorrect pairing is presented (e.g., Object A – Word B). Importantly, the Switch trial consists of a habituated object and a habituated word, but linked in a novel way. In this manner, infants should only dishabituate if they have appropriately linked the word and object.
Methodological Structure

Now that we have discussed the nature of the work using habituation tasks, we can turn to the typical structure of these tasks. Infant habituation studies in language research typically involve four phases: pretest, habituation, test, and posttest. Figure 1.1 outlines these four phases across three different studies. Each of these phases comprises discrete trials wherein a visual and an audio stimulus are concurrently presented. Trials can be preceded by what is termed an attention-getter in order to get the infant to orient to the screen. Various attention-getters have been used in past research. Some examples include: a silent, flashing light; a silent, morphing, colourful shape; and the face of a baby with giggling as the accompanying audio track. Once the infant looks to the screen, the relevant trial commences. Trials can be of fixed length or infant-controlled. The latter involves setting a criterion via which the trial will end if the infant disengages attention. For example, if an infant looks away from the stimulus for 2 seconds, the trial ends and the next trial commences.

**Pretest**

In the pretest phase, the infant experiences a stimulus that is different from the repeated one during the upcoming habituation phase. The reasoning behind this trial is that infants need to become accustomed to the presentation method. Thus, this phase serves as a warm-up prior to presenting the stimulus of import in your study.

**Habituation**

The habituation phase follows and is, of course, key to the experiment. One important point to consider for this phase is the intensity of the audio stimuli. As Thompson and Spencer (1966) highlight in their list of habituation characteristics, “strong stimuli may yield no significant habituation” (p. 19). For example, it would be hard to habituate to a blaring, variable siren. Therefore, the audio stimulus is typically delivered at approximately 65 decibels to make it loud enough for infants to hear, but not too loud to invoke failure to habituate. Following similar logic, the visual display that is shown should also be only moderately engaging. Another setting the researcher must decide upon is the habituation criterion. A common criterion is a 50% reduction in LT (Ashmead & Davis, 1996), although some researchers advocate using more stringent criteria with younger infants (e.g., 70%) as they are more cognitively immature and therefore may require more presentations to fully process stimuli (e.g., Flom & Pick, 2012). Another consideration is the window over which the decrease in response is based. If one stimulus is repeated, many researchers opt for a window of three trials. For example, if the infant looks for a total of 50 seconds over the first three trials, they need to fall below 25 seconds in a subsequent window of three back-to-back trials to reach habituation. Ashmead and Davis recommended the size of this window based on their computer modeling in that it was more stable than a window size of two.
Two other important considerations are related to these windows. First, the researcher could opt for a fixed or a sliding window. A sliding window keeps a running total of LT to determine habituation (e.g., trials 2, 3, and 4 are first compared to trials 1, 2, and 3). A fixed window compares subsequent blocks of three trials to the criterion previous block (e.g., trials 4, 5, and 6 are first compared to trials 1, 2, and 3). Oakes (2010) recommends using the sliding window whenever possible, as it necessarily will lead to shorter experiments on average, and shorter habituation phases should result in less attrition. However, if the infant is being habituated to two types of stimuli (e.g., two word-referent combinations), the fixed window is necessary despite the chance of increased attrition in order to ensure that the infants receive an equal number of examples from each stimulus type during habituation (and the window needs to be increased to four trials: two of each stimuli type per block). The second consideration is whether to base the habituation criterion on the first block of trials, which typically—but not always—has the highest infant behavioral response to the stimuli, or base it on the block of trials with the highest behavioral response, regardless of when it occurs in the experiment. Most researchers use the first block, as infants may have an increased response on a later block due to a factor unrelated to the habituation curve (e.g., a baby surprises himself with a sneeze and reorients to the stimuli due to increased arousal).

Despite the presence of a criterion, the researcher should cap the number of possible trials in an experiment. Without such a cap, the experiment would not end for some infants in a reasonable amount of time, as they would not reach criterion. Dannemiller (1984) recommended that the maximum amount of trials in an infant habituation study should be 15 trials. At that number, according to his Monte Carlo modelling, there would be a 5% risk that infants are habituating by chance. But the trade off of having a small maximum number of trials is that you will have fewer infants reaching criterion within that overall trial limit. But, increasing the maximum amount of trials increases both the chance of random habituation and of attrition. Oakes (2010) recommends piloting infants and/or examining similar studies in the literature to determine the optimal maximum number of trials for a particular study.

**Test**

The test phase of the experiment should include both the novel stimulus and a repetition of the familiar stimulus from habituation, with the order counterbalanced across participants (e.g., Werker et al., 1998). Why not only present the novel stimulus and compare it to the last habituation block? The major issue in taking that approach is that behavioral responses in the final block of habituation trials are necessarily low, and may be artificially so (see Cohen, 2004). This could be due to an infant reducing attention to the stimuli for a reason unrelated to habituation, such as a distraction in the room like the parent shifting in their seat. Thus, this comparison may falsely indicate dishabituation. By presenting both the novel stimulus and a repetition of the familiar stimulus, the experimenter can determine whether the infants can detect the difference between the habituated stimuli and something new. Some researchers run this manipulation as a between-subject design, but this is not recommended for both statistical and practical reasons. One would introduce more error into the design related to the individual differences between the groups, and it requires doubling the amount of participants.
Posttest

Finally, the posttest trial should be presented last and should be maximally different from the habituation and test trials. It is expected that if infants are still engaged in the experiment, looking time would recover to near pretest level during this final trial.

Collecting and Analysing Data

As mentioned earlier, the best method to collect the data is via habituation-specific software, like Habit. The program is designed to continuously compare an infant’s behavioral response to the stimulus, in this case LT, in a block of trials to previous blocks to determine when the infant has reached the habituation criterion. Importantly, before delving into statistical analyses of the habituation and test phases, one has to establish the reliability of the experimenter’s coding of the target variable (e.g., manual key strokes in response to LT) if it is not a direct measure (e.g., eye tracker for LT, ECG for heart rate). One standard is for another coder to recode the target trials of 25% of the participants from the video records of those experiments. In this case, the original coding would be considered to be reliable if a Pearson product-moment correlation of the two coders’ measures is equal to or greater than .95. A more exact method is to have two coders score all the video records using a frame-by-frame analysis using freeware such as SuperCoder (Hollich, 2005).

One should also report analyses of the habituation phase prior to testing for novelty and familiarity effects. To determine whether infants maintained interest throughout the experiment and recovered from habituation, one possibility is to run a series of planned orthogonal comparisons to first compare pretest to posttest and, if these two trials are found to be the same, to then compare these trials to the last habituation block. It is expected that if infants were still engaged in the experiment, looking time would recover to near pretest level during the posttest. Thus, there should be no significant difference between the pretest and posttest. However, the pretest and posttest should be significantly different from the last habituation block to demonstrate recovery. One can then compare the first habituation block to the last via a paired-sample t-test to confirm a significant drop in looking time across the habituation phase. Finally, a full descriptive analysis of the habituation phase should be reported (i.e., mean number of habituation trials, mean looking time during habituation). If there are multiple conditions or groups in the experiment, these habituation variables should be compared in a mixed ANOVA to ensure similar habituation across conditions or groups. For example, in Table 1.1, infants in

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Habituation Trials</th>
<th>Habituation Time</th>
<th>Familiar Trial</th>
<th>Novel Trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>112 s</td>
<td>11.5 s</td>
<td>12.2 s</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>180 s</td>
<td>8.1 s</td>
<td>12.8 s</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>200 s</td>
<td>7.9 s</td>
<td>12.4 s</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>210 s</td>
<td>8.3 s</td>
<td>8.1 s</td>
</tr>
</tbody>
</table>

Note: All data represent the mean.
Experiment 1 and those in Experiment 2 have a similar number of habituation trials, but the latter group has significantly more active looking during habituation. This may explain their success in the task (i.e., significantly higher looking times to novel over familiar stimuli). On the other hand, infants across Experiments 3 and 4 have similar habituation scores, but behave differently at test. Thus, these results cannot be attributed to habituation differences.

To determine if infants have dishabituated to the novel stimulus at test, researchers should compare the novel test trial to the familiar test trial. As test trial is a within-subject variable, a paired-sample t-test is often used. If a between-subjects variable is key to the interpretation of results, a mixed ANOVA would be appropriate. For example, gender is often included in habituation studies testing language skills due to the oft-reported female advantage for language. Indeed, research involving infants’ habituation to word-object associations (i.e., word learning tasks) has found gender differences in terms of a female advantage (e.g., Fennell, Byers-Heinlein, & Werker, 2007). Another common between-subjects variable in infant language research using habituation is age, in order to track developmental changes (e.g., Werker et al., 2002).

Table 1.2 summarizes all steps in collecting and analysing habituation data that were expanded upon in this section.

<table>
<thead>
<tr>
<th>Steps in Temporal Order</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real-Time Habituation Data Collection</td>
<td>To determine if and when habituation occurs, a specialized computer program (e.g., Habit) collects behavioral response data in real time via experimenter input (e.g., button press when infant is looking). It constantly compares blocks of responses to determine habituation.</td>
</tr>
<tr>
<td>Objective Data Collection</td>
<td>Optimally, experimenters should simultaneously collect the relevant behavioral response via a completely objective method (e.g., video recording of infant looking).</td>
</tr>
</tbody>
</table>
| Coding | 1) Confirm the accuracy of the real-time habituation data. Code the raw data for a minimum of 25% of the participants. For looking time, one can use video software (e.g., Super Coder) to examine infants’ looking time in each video frame. Confirm accuracy via a Pearson product-moment correlation of the habituation and coded raw measures that is equal to or greater than .95.  
2) Recommend coding all key trials for all participants via an exact measure (e.g., frame-by-frame coding) to maximize accuracy. |
| Disconfirm Fatigue | Run a planned orthogonal comparison on the pretest, posttest, and last habituation block. The pretest and posttest should be equal to confirm that the baby did not become fatigued over the experiment. The posttest should differ from the last habituation block to confirm recovery to a large change. |
| Confirm Habituation | Compare the first habituation block to the last using a paired-sample t-test to confirm a significant drop in behavioral response. |
| Confirm Dishabituation | Compare the novel test trial to the familiar test trial via a paired-sample t-test. Typically, a significant difference will reveal that the novel trial exceeds the familiar. Rarely, it is the reverse (see discussion in chapter). If a between-subjects variable is key to the interpretation of results, use a mixed ANOVA that includes that factor (e.g., gender). |
To illustrate the use of the habituation task to investigate early language development, a study that involved multiple conditions and populations will be described. Before turning to that study—Fennell and Byers-Heinlein (2014), some background is in order. Using the Switch Task, Byers-Heinlein, Fennell, and Werker (2012) had shown that monolingual and bilingual infants have similar word-object associative skills, reliably noticing the incorrect links in habituated word-object pairings around 14 months when the words involved are phonologically distinct (e.g., /lip/ and /nim/). In contrast to the finding with dissimilar sounding words, Werker, Fennell, Corcoran, and Stager (2002) found that infants growing up in a monolingual environment have difficulty learning similar sounding words in the Switch task up until 17 months of age. Fennell, Byers-Heinlein, and Werker (2007) extended the latter study to bilinguals. Bilinguals of 14, 17, and 20 months received two word-object pairings during habituation: a multi-coloured crown object paired with the nonsense word “bih” and a blue-green molecule object paired with “dih.” Unlike monolingual infants in the previous work, bilingual infants did not dishabituate to the switch in the pairings until 20 months of age (e.g., crown paired with “dih”). Thus, it appeared that bilinguals might be “delayed” in incorporating their emerging phonology into word learning, in that they were accepting a similar-sounding, but incorrect, word as an appropriate label for the object. Fennell et al. argued that ignoring detail might be adaptive for bilinguals (less is more): by incorporating less information into the word form, they could match their monolingual peers on word learning.

However, a subsequent study by Mattock, Polka, Rvachew, and Krehm (2010) indicated that there was more to the story. Fennell et al. had used recordings of a monolingual English speaker in their study. Mattock and her colleagues showed that French-English bilingual infants of 17 months succeeded if given a mix of French and English tokens produced by a French-English bilingual speaker. Monolingual infants failed with these mixed stimuli. Thus, they discovered inverse results in comparison to Fennell et al.: bilingual success and monolingual failure. Mattock et al. argued that bilinguals may have enhanced flexibility in their phonological representations, allowing for success with the mixed tokens in the face of monolinguals’ failures. However, Mattock et al. did not test bilingual infants on monolingual-like stimuli.

Fennell and Byers-Heinlein (2014) hypothesized that both previous interpretations may be incorrect. There may be no bilingual delay or advantage in learning similar-sounding words. Instead, it may be that both monolinguals and bilinguals may simply have difficulty with a “non-native” speaker, as Fennell et al. used a monolingual speaker and Mattock et al. used a bilingual one. Bilingual and monolingual speakers have differing productions, even in the same language (Antoniou, Best, Tyler, & Kroos, 2011). Therefore, this hypothesis would be in line with recent work showing that infants have difficulty with accented speech (Schmale, Hollich, & Seidl, 2011). To determine if our hypothesis was correct, we implemented a fully crossed design where English monolingual and English-French bilingual infants were tested on monolingual- and bilingual-produced target words.

Infants were tested in the standard Switch task (see Figure 1.1). All trials were fixed at 20 seconds each (i.e., not infant controlled). The first trial was a pretest trial
Habituation Techniques

consisting of a word-object pairing (a spinning waterwheel toy paired with /nib/) completely dissimilar to those presented in the habituation phase. The two habituation pairings were the crown object paired with /kɛm/ and the molecule object paired with /gɛm/. Each pairing was presented twice within a four-trial block, with no more than three consecutive trials of the same type. When average looking time across a four-trial block decreased to 65% of the maximum looking time across a previous block, the habituation phase ended. There was a maximum of 24 habituation trials—the typical amount for the Switch task. During the test phase, infants received one Same and one Switch trial in one of eight testing orders that counterbalanced trial order (Same-Switch/Switch-Same) and the particular pairings presented. The posttest was a repetition of the very dissimilar stimuli presented during the pretest. Infants sat on their parents’ laps during testing. All video stimuli appeared on a screen directly in front of the infant and audio stimuli were presented via speakers below the screen at approximately 65 db. Infant looks were recorded via a hidden video camera below the screen. Videos of all test trials were coded frame-by-frame (i.e., each video frame was examined to determine if infants were looking to the target) by an experienced coder, blind to condition, and then a second coder recoded 25% of those trials to ensure high reliability.

Sixty-one infants of 17 months successfully completed the study: 31 monolingual and 30 bilingual. Sixteen infants in each group heard tokens of the target words produced by a bilingual speaker and 14 infants in each group heard tokens produced by a monolingual speaker. An additional 24 infants were tested but not included in the analyses due to fussiness, parental interference, distraction during testing, or being off-camera during the test trials. This is a normal attrition rate for this age group.

The habituation and test trial data are presented in Figure 1.2. The first analysis confirmed that infants across all conditions habituated to the stimuli. A 2 (habituation block: first versus last) × 2 (token: monolingual or bilingual) × 2 (infant: monolingual or bilingual) ANOVA revealed a significant decrease in looking time in the last habituation block compared to the first in all conditions, with no interactions. The number of trials and total looking time it took to reach the criterion across conditions was statistically equivalent. Thus, any differences at test could not be attributed to differences in exposure during habituation. All groups of infants also showed significant recovery during the posttest as compared to the last habituation block. Thus, infants were not fatigued or generally disinterested in the task. Since all habituation measures were normal, an analysis of the test trials could be undertaken.

The key parts of the test trial analysis were the test trials themselves, whether the stimuli matched the infant’s language environment (stimuli match; e.g., a monolingual infant hearing monolingual tokens), and the language background of the infant. A 2 (trial type: Same versus Switch) × 2 (stimuli match: yes versus no) × 2 (infant language background: bilingual versus monolingual) mixed ANOVA revealed a significant main effect of trial type, moderated by a significant interaction between trial type and stimuli match. No other effects were found. Thus, infants’ comparative looking across the Same and Switch trials depended on whether the tokens they were hearing matched their learning environment, and monolinguals and bilinguals showed the same pattern since there was no effect of language background. Follow-up t-tests comparing the Same and Switch trials in each condition, with the appropriate corrections for multiple tests, revealed that infants only detected the novel pairing (i.e., a change in minimally different labels) if they heard tokens that matched their
Research Methods in Psycholinguistics and the Neurobiology of Language

language background: bilinguals hearing bilingual tokens and monolinguals hearing monolingual tokens. Thus, the relatively simple habituation task revealed that bilingual infants are neither advantaged nor disadvantaged relative to monolinguals, and helped to clarify the phonological development of all infants by showing that the optimal stimuli for learning words are those presented in a manner similar to infants’ everyday language environment.

Problems and Pitfalls/Advantages and Disadvantages

Some of the advantages of habituation tasks have already been highlighted. The reliability of the method is a key advantage, especially when dealing with a child population. Multiple studies have demonstrated that infants’ habituation responses are stable in the short- (e.g., Bornstein & Benasich, 1986) and long-term (e.g., Miller et al., 1979). As Bornstein and Benasich emphasized, this psychometric reliability gives strong validity to the use of habituation as a methodological tool in infant research (p. 97). The simplicity of the method and the relatively low technology needs make such tasks very easy to implement. For example, if one were testing phoneme discrimination, the use of a habituation technique would involve much less cost than implementing an ERP study, another valid method for determining discrimination skills. This cost imbalance would be both in terms of the technology needed, with ERP equipment (multiple computers, software, EEG caps, etc.) running...
in the thousands of dollars, and in terms of the time involved (training, coding, analyzing, etc.). Habituation tasks also have a long history both inside and outside of their use with infants. This provides the researcher with a large literature base from which to design and explore relevant tasks, and to interpret possible results. Newer methods, such as fNIRS, do not yet have a rich literature, leading to more guesswork.

Another advantage of habituation is its possible use across all ages. For example, the Conditioned Head Turn procedure, another method that has often been used to determine infants’ phoneme discrimination skills, has a limited age range. In this technique, infants are trained to turn their head upon detecting a phoneme change. Werker, Polka, and Pegg (1997) suggest that the procedure is optimal for a small developmental window of between 6 and 10 months, as infants below 6 months have limited head control and older infants become too “mobile” to sit through the long task involving training and testing. Habituation can be used throughout the lifespan, from foetus to adult. It is a basic psychological process and is tailored to the individual’s cognitive skills. However, it should be noted that infants do become more restless in a habituation task as they become older and more mobile. The application of a habituation criterion appropriate to the infant’s age (based on similar past studies with said age) and close measurement of their responses should alleviate this issue.

Another technique often used in infant language studies is preferential looking (e.g., Looking-While-Listening, Intermodal Preferential Looking Procedure; see Chapter 2). These tasks often involve presenting two visual stimuli simultaneously to a child and measuring looking to each while hearing an auditory stimulus associated with one of the visual displays. Often, researchers use this method to test word knowledge (e.g., infant sees a dog and a cat on a screen and looks longer to the dog when hearing “dog”). Another variant is the Head Turn Preference procedure (see Chapter 2) where the infant turns her head toward a visual stimulus to hear an auditory stimulus, which ends when she turns away. Infants’ preference for one stimulus over another (e.g., native versus non-native language) is inferred from the amount of time they listen to each. The above examples reveal that these techniques are often used to test infants’ preference for stimuli learned in their natural language environment. Habituation, on the other hand, necessarily teaches the infant new information (e.g., new words) or re-familiarizes them to information from their environment (e.g., language sounds), always to a criterial level. As such, the advantage of habituation is the ability to ensure that infants have processed the stimuli immediately prior to testing. A comparative disadvantage is that habituation involves more memory load at test: it presents test trials sequentially, whereas preferential looking presents choices simultaneously so that all information is available. Hybrid tasks that use a habituation phase followed by a visual preference phase maximize the benefits of both techniques (Yoshida, Fennell, Swingley, & Werker, 2009). One large advantage of visual preference tasks is that a researcher can present multiple stimuli during the course of an experiment (e.g., test knowledge of multiple words), as no training phase is required. Habituation is limited to comparing a familiar versus a novel stimulus.

A technique similar to habituation is familiarization. Studies that present stimuli until infants reach a certain LT amount (e.g., 5 minutes total) or a certain amount of trials are not habituation experiments, but rather are familiarization studies.
These studies suffer from a major issue: Unlike habituation, they fail to tailor the learning phase to the individual participant. All infants do not require the same amount of time to learn about a stimulus. Some may take 30 seconds to process the information presented, while others may take 2 minutes. By failing to control for individual learning via a criterion, these studies have a greater chance of producing strange effects, such as a preference for a familiar stimulus over a novel stimulus (which we will turn to next) or producing null results since a segment of the participants may not have processed the stimuli.

An apparent problem one can encounter when using the habituation task is the presence of what is called a familiarity effect. After habituation to a stimulus, infants may demonstrate a reduction of a behavioral response to a novel stimulus, but an increase or maintenance of the response to a repeated presentation of the habituation stimulus. This runs counter to the novelty effect typically found in habituation: increased behavioral response to a novel stimulus and decreased or maintained response to a repetition of the habituation stimulus. However, this atypical response may not be a real problem at all, but rather an informative reflection of infants’ processing of the stimuli.

Familiarity and novelty effects are tied to the level of difficulty infants encounter in processing habituation stimuli. Infants avoid or show no interest in stimuli that are not at their optimal level of stimulation (Cohen, 2004; Hunter & Ames, 1988). Since infants are cognitively immature, they initially actively avoid a complex stimulus and only start preferring it as they become familiar with its properties. Infants exposed to too complex information during habituation may show a “preference” for the familiar stimulus at test, as they are still trying to process its components and actively reject new complex information (i.e., the novel stimulus). For example, if we return to Figure 1.2, the looking time to the Same trial in the condition where monolinguals were hearing tokens from bilingual speakers appears to be higher than expected based on the data from the other conditions. This difference may be reflective of a weak familiarity effect due to increased complexity in that one condition. Infants being raised in a monolingual environment may have had few opportunities to hear bilingual speakers, whereas bilingual infants often hear monolingual speakers. Of course, the other two conditions were also less difficult because infants were hearing typical voices from their environment.

It should be noted that Cohen (2004) argues that we should adopt strict habituation criteria, such as a 50% reduction in a behavioral response, to avoid possible familiarity effects. Of course, the use of a strict criterion can lead to the problem of increased attrition as infants can cross the line from bored to extremely restless and non-participatory. An examination of past similar research with infants of the same age should give researchers solid ideas of both an appropriate habituation criterion and what stimuli are optimal for infants that age.

Another issue is that infants can have preferences for some habituation stimuli, which would interfere with or prevent habituation, or can lead to false familiarity or novelty preferences at test (Oakes, 2010). Again, the appropriate choice of a habituation stimulus is key: Do not use stimuli for which infants have strong preferences and ensure that the familiar and novel stimuli have similar preference strengths in the population being tested. These preferences can be determined via pilot testing, or from a literature review of the target area.

Finally, one cannot automatically treat habituators and non-habituators as the same population. Non-habituators are those infants who reach the maximum
number of habituation trials without meeting the criterion of a reduced behavioral response to a set percentage. Cohen (2004) highlights that, in comparison to habituators, there is a greater chance that non-habituators will alter the results of a study by demonstrating a familiarity preference at test, due to their incomplete processing of the stimulus. For example, Werker et al. (1998) found that, at 14 months of age, only habituators showed evidence of learning a word-object pairing at test, whereas non-habituators did not. Having highlighted these differences, it is important to note that some studies have found no group differences between habituators and non-habituators (e.g., Byers-Heinlein, Fennell, & Werker, 2012). However, studies should always compare habituators and non-habituators to determine if any differences are present.

In conclusion, the simplicity of the habituation procedure in both its design and implementation, along with its long history in our field, makes this task one of the fundamental tools that psycholinguists can use to uncover nascent, emerging, and maturing language skills during infancy and early childhood.

**Key Terms**

**Familiarity preference** An uncommon response in a true habituation task where the participant attends more to the familiar (i.e., habituation) stimulus at test than to a novel stimulus. This usually would indicate that the wrong habituation criterion was employed and/or that the familiar stimulus was too complex for the participant.

**Familiarization study** Unlike a habituation experiment with its individualized criterion, this is a study where every participant experiences the target stimuli for the same predetermined amount of time.

**Familiar stimulus** The stimulus that the participants receive throughout habituation.

**Habituation** The progressive reduction of an organism’s behavior in response to a repeated stimulus.

**Habituation criterion** The set percentage to which the participant’s behavioral response must decrease from the maximum response during habituation (sometimes the maximum during the first block of trials only) before the test phase begins.

**Habituation curve** The pattern of the participant’s responses over the habituation phase, typically an exponential decrease in response—thus the term “curve.”

**Novel stimulus** A test stimulus distinct from the familiar, or habituated, stimulus. Participants should increase their target response to this stimulus over the familiar one, if they have adapted to or learned the latter stimulus from habituation (see Novelty preference).

**Novelty preference** The classic test response in a habituation task where the participant attends more to a novel stimulus than to a familiar one post-habituation.

**Switch procedure** An associative word-learning variant of the habituation task where participants receive two word-object associations throughout habituation (Object A – Word A; Object B – Word B) and are tested on two test trials, a familiar pairing (Object A – Word A) and a novel one (Object A – Word B). If the participants learned the associative link, they will show a novelty response.
References


Further Reading


