PSYCHOACOUSTICS-WEB

Introduction

Adaptive procedures are the most efficient techniques to measure sensory thresholds (e.g., see Watson and Fitzhugh, 1990). These techniques belong to psychophysics since its origin (Fechner, 1889) and are still developed up-to-date (e.g., Faes et al., 2007). However, they are not used as much as one would expect. The reason for this may be the lack of available research tools to implement these procedures in a simple and effective manner. Common licensed softwares for stimulus presentation (e.g., E-prime, SuperLab, etc.) often do not offer these procedures. In contrast, free research tools do but require certain programming skills of the user (e.g., PsychToolbox, Brainard, 1997; Pelli, 1997; APEX, Geurts and Wouters, 2000; APEX 3, Francart et al., 2008).
In more recent years, Grassi and Soranzo (2009) and Soranzo and Grassi (2014) implemented a MATLAB toolbox running several classic adaptive algorithms for threshold estimation. Although these toolboxes fulfilled the needs of several researchers for several years, many users are still intimidated by the software per se, by the installation of the toolbox, not to mention the frequent updates of the MATLAB software that are often incompatible with the characteristics of the scripts written for the toolbox. Last but not least, MATLAB is a quite expensive software. Therefore, it is often not an option for research groups that lack of fundings. Nowadays, tools like those implemented by Grassi and Soranzo (2009) and Soranzo and Grassi (2014) can fit easily in a website and PSYCHOACOUSTICS-WEB is the direct translation of the PSYCHOACOUSTICS into a web site. In the remaining text, readers will find some introduction to threshold estimation and instructions on how to use the toolbox.

README FIRST

The current version of PSYCHOACOSTICS-WEB implements only few algorithms for threshold estimation (i.e., those of the staircase family) and implements only three tasks: pure tone frequency discrimination, pure tone duration discrimination, and pure tone intensity discrimination. Future versions will implement further tasks (they will be listed together with the present ones) and further adaptive algorithms.

Threshold estimation

Sensation moves within and across two types of thresholds: detection and discrimination. The detection threshold is the minimum detectable stimulus level in the absence of any other stimuli of the same sort. In other words, the detection threshold marks the beginning of the sensation of a given stimulus. The discrimination threshold is the minimum detectable difference between two stimuli levels. Therefore, for a given sensory continuum, the discrimination threshold cuts the steps into which the sensory continuum is divided.
The detection threshold can be estimated either via yes/no tasks or via multiple Alternative Forced Choice tasks (in brief nAFC, with n being the number of alternatives). The

discrimination threshold, in contrast, is estimated mainly via nAFC tasks. In yes/no tasks, the subject is presented with a succession of different stimuli levels (spanning from below to above subject's detection threshold) and is asked to report whether s/he has detected the stimulus (yes) or not (no). In nAFC tasks, the subject is presented with a series of n stimuli differing in level. In audition, because the various stimuli have to be presented in temporal succession, tasks are often multiple intervals tasks (i.e., mI-nAFC). In a nAFC task, one stimulus (the variable) changes its level across the trials; whereas the level of the others (the standards) is fixed. The difference between standard and variables ranges from below to above subject's detection (or discrimination) threshold. After each trial, the subject is asked to report which the variable stimulus was. Figure 1 shows the hypothetical results of a yes/no task.

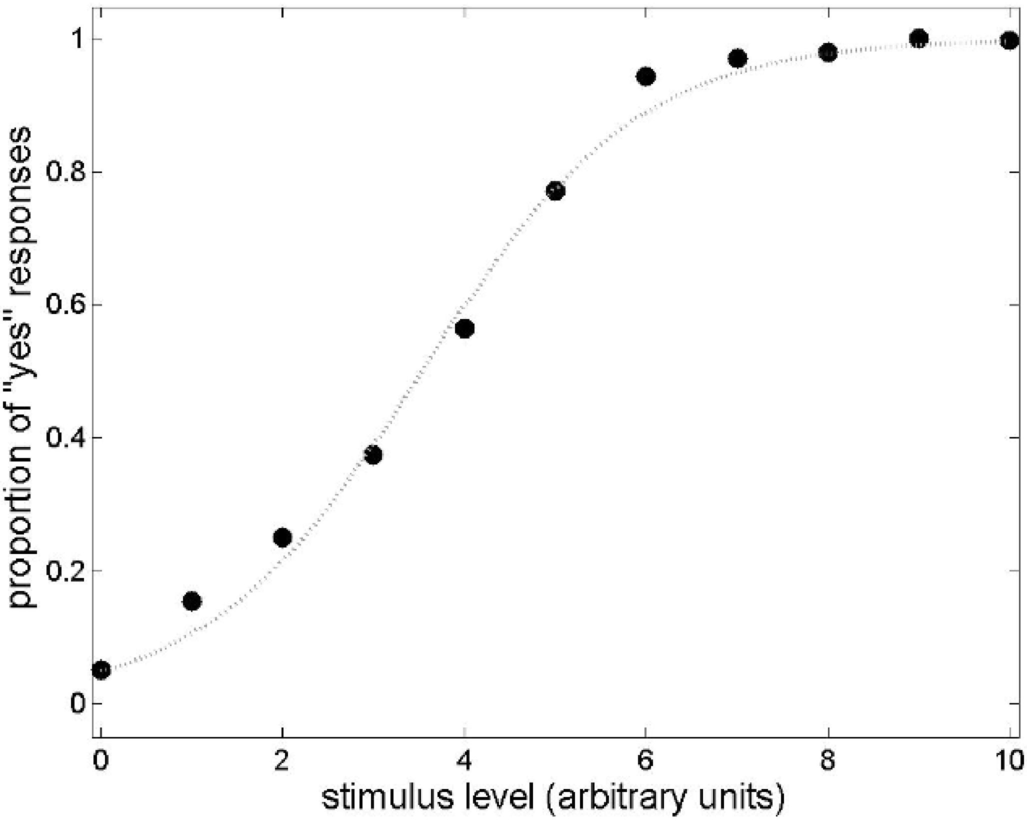

Figure 1. Results of a hypothetical yes/no task. Subject's data are fitted with a logistic function (dashed curve). Note that this subject committed some false alarms (see later in the text) because at zero stimulus level (i.e., no stimulus actually presented) we can still observe a certain number of yes responses.

Figure 1 shows the relation between the stimulus level and the subject's performance together with one function fitting the hypothetical data. This function is referred to as the psychometric function. Independently from the task type, and from the type of threshold being measured, behavioural data are hypothesised to fit with a sigmoid function such as that represented in Figure 1. Different types of psychometric functions can be adopted to fit experimental data, for example, the logistic, the Weibull and the cumulative Gaussian.
The general equation of a psychometric function is the following (adapted from Wichmann & Hill, 2001) representing the subject's performance as a function of the stimulus level x.

$$\Psi(x;\ \alpha, \beta, \gamma, \lambda) = \gamma + (1 - \lambda - \gamma) f(x;\ \alpha, \beta)$$

f(x) is the sigmoid function chosen by the experimenter (i.e., logistic, Weibull, cumulative Gaussian). β determines the function's slope, whereas α determines the displacement of the function along the abscissa (see Figure 2). γ and λ are "psychological" parameters and they will be discussed shortly.

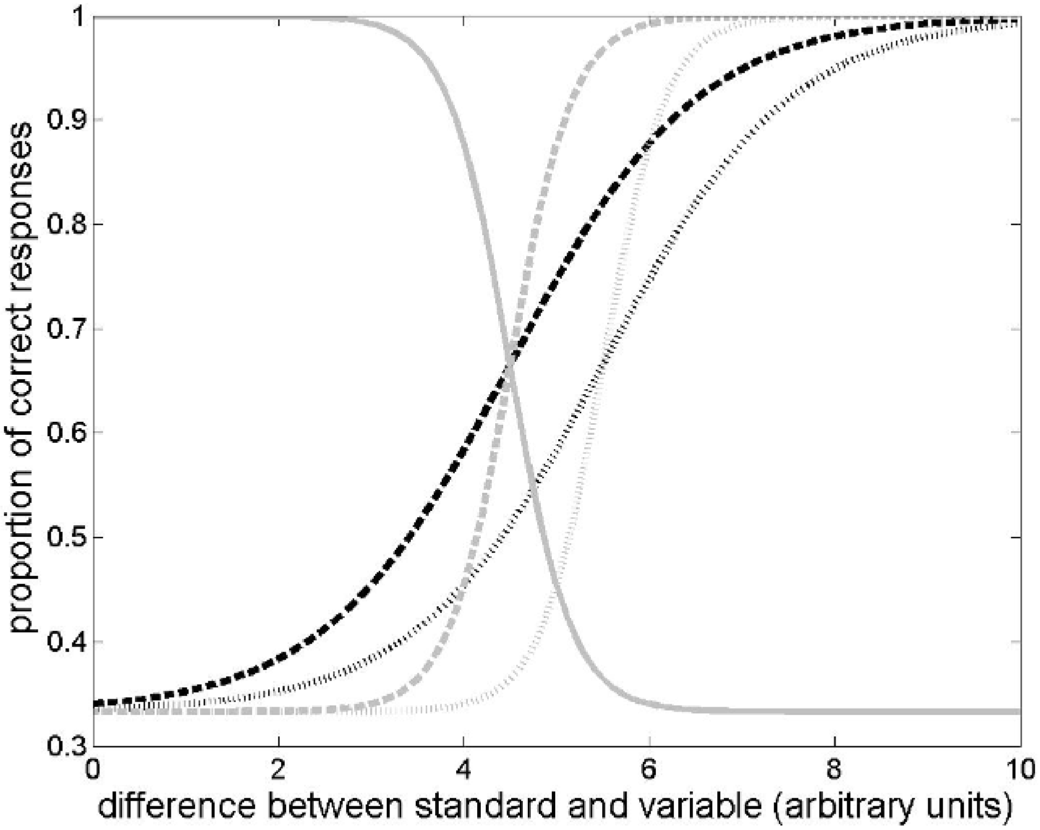

Figure 2. Five logistic psychometric functions. The grey (or black) functions have identical slope. The dashed (or dotted) functions have identical midpoint. The grey, solid function has a negative slope.

Amongst the sigmoid functions, the logistic is the most widely used, because of its computational simplicity. Its formula is the following:

$$f(x) = \frac{1}{1+e^{\beta(\alpha - x)}} \qquad [2]$$

Therefore, the corresponding psychometric function Ψ is:

$$\Psi = \gamma + (1 - \lambda - \gamma)\left[\frac{1}{1+e^{\beta(\alpha - x)}}\right] \qquad [3]$$

In the logistic psychometric function, α (often referred to as midpoint) enables the displacement of the function along the stimulus level axis. It corresponds to the average between γ and λ [i.e., $\Psi(\alpha)=(\gamma+\lambda)/2$]. β is the function slope, i.e., the rate of change in the subject's performance with stimulus level. The greater the absolute value of β, the steeper the psychometric function will be. Moreover, for positive values of β the function increases whereas for negative values it decreases (see Figure 2). γ and λ come into play when adapting the function to psychological needs. γ assumes a different meaning depending on the task type (i.e., yes/no or nAFC). There is, in fact, a major difference between these tasks: in yes/no tasks, the subject's response criterion is not under control of the experimenter, on the contrary, it is in nAFC tasks (Green & Swets, 1966; Stanislaw & Todorov, 1999). The reason for this difference is that in a yes/no task, when a yes response is collected for a very low stimulus level, it is difficult attributing this response to a high subject sensitivity or to a bias toward the yes response. Biased responses are called false alarms and they affect the lower limit of the psychometric function, which can assume values greater than zero (see Figure 1). In other words, the probability to get a yes response in absence of the stimulus is greater than zero (Green, 1993). Hence, in yes/no tasks, γ corresponds to the subjects' false alarm rate. False alarms are absent nAFC tasks (Green & Swets, 1966; Stanislaw & Todorov, 1999). In nAFC tasks, the level of the standard(s) is always different from the level of the variable and trials have, therefore, correct and incorrect answers. When the difference in level between standard and variable is below the subject's threshold (i.e., "when the stimulus level is low") the probability that the subject returns a correct answer is determined by chance, and chance will depend on the number of alternatives. For this reason, in nAFC tasks, γ corresponds to chance level, i.e., the reciprocal of the number of alternatives (e.g., 50% for 2AFC, 33% for 3AFC, 25% for 4AFC and so on).

The meaning of λ, on the contrary, is independent of the task and refers to another error. In both yes/no and nAFC tasks, subjects could commit errors independent from the stimulus level; they are the lapses of attention (i.e., "I want to respond "A" but I erroneously press the response key "B"). Lapses of attention are estimated to be a small percentage of the subject's responses (i.e., 1-5% Saberi & Green, 1997; Wichmann & Hill, 2001) and they can affect the psychometric function fitting by decreasing the upper limit of the function (see Wichmann & Hill, 2001 for an extended discussion). Attentional lapses are particularly problematic at high stimuli levels - in yes/no tasks - or high differences in the level of between the standard and variable(s) - in nAFC tasks (Saberi & Green, 1997; Wichmann & Hill, 2001).

Researchers are often interested in estimating a single point of the psychometric function, which is the subject's threshold. In probabilistic terms, the threshold corresponds to an arbitrary point of the psychometric function p (hereafter referred as to p-target) included between the lower and the upper limit of the function (i.e., γ and λ). In other words, when we estimate a threshold, we search for the stimulus level eliciting the p-target proportion of yes (or correct) responses. Treutwein (1995) proposes that the p-target should be the middle of the psychometric function (e.g, 50% for yes/no tasks, 75% for 2AFC, 66% for 3AFC, etc.). However, other authors suggest that higher values should be targeted (e.g., Green, 1990; Baker & Rosen 1998, 2001; Amitay, Irwin, Hawkey, Cowan, & Moore, 2006).

Thresholds can be estimated by means of two classes of procedures: adaptive and non adaptive. In non-adaptive procedures, for example the constant stimuli method, the stimuli levels (or differences between standard and variable level) are preset before the beginning of the experiment. The stimuli should span from below to above the subject's threshold. During the experiment, the stimuli are presented to the subject in random order and the proportion of yes (or correct) responses is calculated for each stimulus. In other words, the subject’s threshold will be interpolated from a fully-sampled psychometric function making the measurement of the threshold expensive in terms of experiment’s time. This represents the major drawback of this class of procedures when the experimenter needs to estimate the subject’s threshold only. Another drawback is that often the experimenter does know where the threshold may be. Therefore, s/he may sample badly the stimuli for the experiment. For the above reasons, when they need to estimate a threshold, psychophysicists prefer adaptive over non adaptive procedures. In adaptive procedures, the stimuli levels are selected at the same time as the experiment is running, depending on the subject answers. Adaptive procedures maximise the ratio between number of stimuli presented at/near threshold and number of stimuli presented far from threshold. Adaptive procedures can be grossly divided in two types: nonparametric (also known as staircases) and parametric. The only assumption made by non-parametric procedures is that the psychometric function is monotonic. Parametric procedures, on the contrary, make more assumptions. For example, they assume the shape of the psychometric function. Examples of non-parametric procedures are the method of limits by Fechner (1889), the simple up-down by von Békésy (1947) and the transformed up-down by Levitt (1971). Examples of parametric procedures are the PEST, by Taylor and Creelman, (1967), the “best” PEST by Pentland (1980) and the QUEST by Watson and Pelli, (1983), and the maximum likelihood (Green 1990, 1993). Nonparametric procedures are generally more used than parametric ones, even if they encompass some disadvantage. The major one is that they tend to be more time consuming (e.g., Amitay et al., 2006; Leek, 2001). However, non parametric procedures are theoretically simpler and can be easily implemented via conventional software (e.g., E-Prime), whilst parametric procedures are theoretically more complex and require more advanced programming skills.

The “staircase” algorithm

Staircases are perhaps the oldest adaptive psychophysical procedures. Three variants can be distinguished within this type of procedure: the method of limits (Fechner, 1889), the simple-up down (von Békésy, 1947; Dixon and Mood, 1948) and the transformed up-down (Wetherill and Levitt, 1965; Levitt, 1971). The staircases can be used to estimate both detection and discrimination thresholds. Moreover, they can be used with yes/no tasks and with nAFC tasks. Because staircases have been used for decades, they have been extensively investigated (e.g., Garcia-Perez, 1998, 2002, 2009) also in comparison with more contemporary procedures (for example, in auditory research, Amitay et al., 2006; Kollmeier et al., 1988; Marvit et al., 2003). These investigations reveal that the staircases are still a very reliable way to estimate sensory thresholds. In some cases, authors also suggested possible improvements for staircases (e.g., Brown, 1996; Garcia-Perez, 2009). However, these improvements did not find practical application nor had a real impact in everyday laboratory use. Therefore, here staircases are described in their most commonly used variants and implemented in the most classic way in PSYCHOACOUSTICS-WEB.

The method of limits

The method of limits is commonly attributed to Fechner (1889) although this attribution has been questioned by Boring (1961). To describe this method, let us consider the case in which we wish to estimate the frequency discrimination threshold of a 1-kHz pure tone. There will be two types of stimuli: the standard and the variable. The frequency of the

standard is fixed. The frequency of the variable is higher than the standard of a certain value Δf, and the value of Δf is adaptively changed during the experiment as a function of the subject’s response. In each trial, standard and variable are presented in random order and the subject is asked to report the tone with the highest pitch. Every time the subject's answer is correct, Δf is reduced. At a certain trial n, the subject's answer will be incorrect because Δf is below threshold (and because the subject has not guessed the right answer). This is a so called “reversal” because from a series of correct answers we are now registering an incorrect answer. The reversal is also the point where we crossed the subject’s threshold because we passed from the last positive response, to the first negative response. The threshold corresponds to the average between Δfn and the Δfn-1; that is, the average between the stimulus level before and after the reversal (Figure 3, left graph, trial 8-9).
By means of this calculation, the method of limits returns the stimulus level corresponding to 50% of the psychometric function. In fact, threshold calculation is made with the last level returning a correct answer (i.e., 100% of the psychometric function) and with the first level returning an incorrect answer (i.e., 0% of the psychometric function). Theoretically, the method of limits (but also the simple and the transformed up-down) can also be run in the opposite order; the first level of the variable can be below the expected threshold and in the subsequent trials the level of the variable is increased instead of decreased. However, in psychoacoustics, the threshold is rarely approached from below.
If the initial values of Δf and Δf changes are carefully selected, the method of limits is the fastest and simplest method in psychophysics to approach the threshold. However, the rapidity of the method is overtaken by the influence of chance in nAFC tasks and the influence of false alarms in yes/no tasks (Gescheider, 2003). For these reasons, the method of limits is nowadays rarely-to-never used.

The simple up-down

To solve some of the drawbacks of the method of limits, Nobel Prize von Békésy (1947) proposed the simple up-down (but see also Dixon and Mood, 1948). The simple up-down does not stop tracking the threshold at the first reversal: when the first reversal occurs, the procedure continues tracking the threshold until a pre-set number of reversals occurs. To see how this works, let us consider again the frequency discrimination example. When the subject returns the correct choice, Δf is reduced; and when the subject returns an incorrect response, we record the first reversal. However, in contrast with the method of limits, the experiment does not stop here and the subject is presented with at least one more stimulus having an increased Δf. For example, we could again present the same stimulus presented before the reversal (Figure 1, right graph, trial 9-10). To summarize, every time the response is correct Δf is reduced; whilst every time the answer is incorrect Δf is increased. Similarly to the method of limits, the simple up-down method also tracks 50% of the psychometric function. Note that the method of limits is actually a particular case of simple up-down: it is a simple up-down in which we stop at the first reversal.

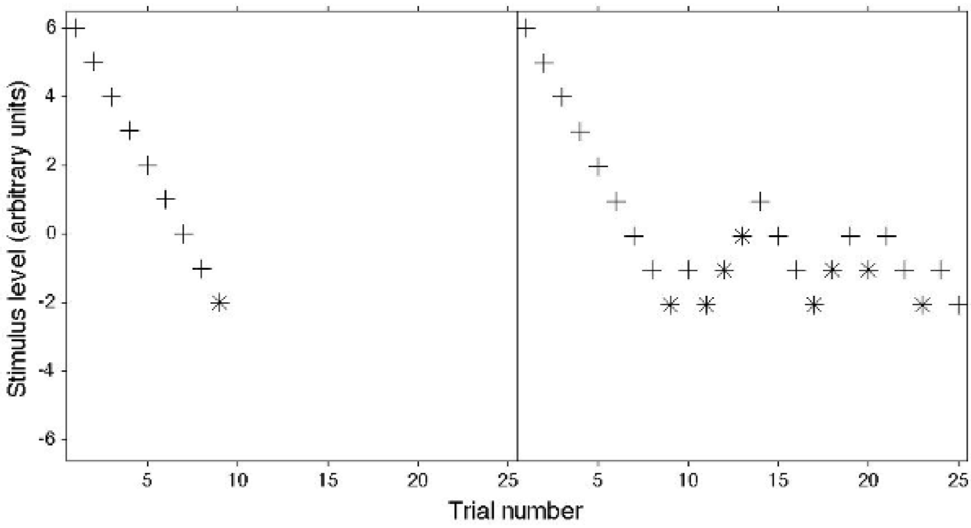

Figure 3. Hypothetical threshold tracking with the method of limits (left) and with the simple up-down (right). The plus sign represents the correct responses whereas asterisk represents the incorrect responses. Note that the threshold trackings are identical up to trial n. 9. Both trackings start with a stimulus level of 6 and reduce the level of step size by 1 at each reversal.

The transformed up-down

The transformed up-down (Wetherill and Levitt, 1965; Levitt 1971), can track different points of the psychometric function. This is because the up and down change of the threshold tracking is unbalanced. Indeed, in both the method of limits and in the simple up-down, change in the threshold tracking is balanced; that is, the variable stimulus goes toward threshold after one correct response and moves away from threshold after one incorrect response. For this reason the simple up-down is also defined as 1-up, 1-down procedure. In the transformed up-down, the variable stimulus moves down, toward threshold, for two (or more) positive responses, whilst it moves up after one negative response only.
Let us suppose the probability of a stimulus giving rise to a positive response is *p*. In this case, the transformed up-down suggests moving down when the subject returns two or more positive responses and to move up when the subject produces one negative response. Therefore, the probability of moving down, toward threshold, becomes p^2 whereas the probability of moving up, away from threshold, is either 1-*p* (i.e., one negative response only) or *p*(1-*p*), i.e., one positive response followed by one negative response. In synthesis:

$$p^2 = (1 - p) + p(1 - p) = 1 - p^2 = 0.707$$

It is for this reason that the so-called 2-down 1-up method tracks 70.7% of the psychometric function.
There are many possible variants of this method. The most popular, together with the 2-down 1-up, is 3-down 1-up which tracks 79.4% of the psychometric function (i.e., =.794). It should be noted that each time the number of responses moving down is increased (e.g., from 2-down to 3-down), the length of the experiment increases because each group of “down” responses is lengthened by at least one trial. Transformed up-down are perhaps the

most common procedures used in psychophysics. They become particularly interesting when the algorithm is used in a forced choice task. For example, if we adopt a 2AFC task the minimum performance we have to expect is 50%, because when the stimulus level is extremely low, the subject can still guess the correct response by chance. However, because transformed up-down track higher points of the psychometric function (e.g., 70.7% or 79.4%) the researcher can still collect a good sensory measure that is not much affected by chance.

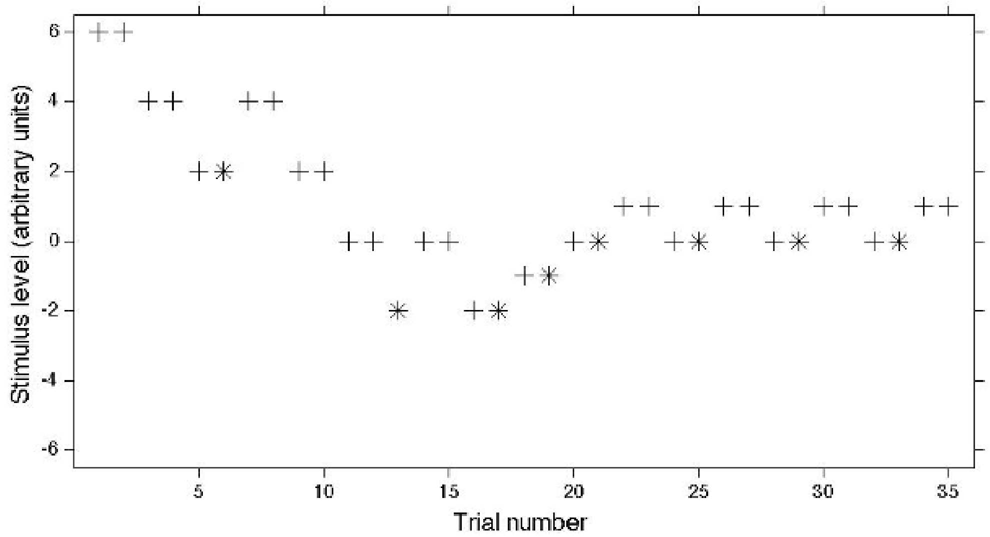

Figure 4. Hypothetical threshold tracking with the transformed up-down. The plus sign represents the correct responses whereas asterisk represents the incorrect responses. The starting stimulus level is 6. The total number of reversals is twelve. The first four reversals are performed with a step size of 1 and the successive eight are performed with a step size of 0.5. Note how the transformed rule lengthens the threshold tracking in comparison with the method of limits or the simple up-down procedure (see Figure 3).

How to change the stimulus level.

If we use a staircase there is one main way by which we can change the stimulus level during the threshold tracking: by multiplication/division. We call “factor” the factor by which we multiply (or divide) the value of the current delta when we need to change its value. For example, let us suppose we are estimating the frequency discrimination threshold for a 1-kHz pure tone. We may set a standard tone of 1 kHz and a variable tone of 1-kHz plus a certain delta *f*. Let's suppose to set a factor of 2. If we are approaching the threshold from above, every time we observe a positive response (or a group of responses that forms a positive response) we have to divide delta by the factor. On the contrary, if we observe a negative response (or a group of responses that forms a negative response) we have to multiply delta by the same factor. In laboratory practice it may be convenient to use more than one factor. It is common in psychoacoustics to use two factors: a large one (e.g., 2), to approach the threshold quickly and a smaller one (e.g. $\sqrt{2}$), for fine threshold estimation. A common solution is to adopt the large step size in the first 4 reversals and the small one in the last 8 (or 12) reversals. In any case, factors should not be chosen to be excessively large or excessively small: a too large step size would produce an alternance of very easy trials and very difficult trials; a too small factor increases the accuracy of the threshold measure but at the cost of lengthening the experiment.

How to calculate the threshold.

With the method of limits the threshold is equal to the average between the last two deltas before and after the reversal. The threshold calculation is slightly different in simple and transformed up-down. In both procedures, the threshold tracking is divided in so-called "runs". One run is a set of consecutive trials including one reversal at the end. Because each reversal is a threshold estimate, the simple up-down and the transformed up-down gather several threshold estimations. Usually, the threshold is calculated by averaging the various thresholds collected during the runs. Figure 1 shows a possible threshold tracking arising from the simple up-down staircase. In the case shown in Figure 1, reversals occurred at trials 8-9, 9-10, 10-11, 13-14, 16-17, 18-19, 19-20, 20-21, 22-23, and 23-24. We may decide to calculate the average of the thresholds at the last two reversals (e.g., stimuli levels –0.5 and –1.5 in the example of Figure 1). In everyday lab-practice experimenters tend to discharge (at least) the first reversals of the threshold tracking and calculate the threshold only on the latest reversals. This is particularly true when the first reversals are obtained with a large factor (or step size). In conclusion, in the case of the simple and the transformed up-down, the threshold is calculated by averaging (either arithmetically or geometrically) the various thresholds at the reversal points. Alternatively, the median can also be used.

Deciding which staircase is good for you: guidelines

In the current section we describe a few guidelines that may help the user to set as efficiently as possible the staircase's characteristics for her/his needs. First of all, the reader needs to be aware that there is a general trade off in psychophysics: robust threshold estimation requires long-duration experiments. For example, the duration of nAFC experiments increases with the number of alternatives. In audition, nAFC tasks are unlikely to exceed the three alternatives otherwise the experiment duration becomes excessive (Schlauch and Rose, 1990). However, the higher is the number of alternatives the smaller is the effect of chance in the threshold estimation and consequently the better is the threshold estimation. The time/accuracy principle applies also to the variants of the staircase family. Passing from the method of limits to the simple up-down the experiment duration increases because of the increased number of reversals. Moreover, moving from the simple up-down to the transformed up-down the duration of the experiment increases as a function of the number of "downs". The transformed up-down in particular reaches a good compromise between duration and accuracy when combining the 2-down, 1-up tracking option with a 3AFC task or, in alternative, the 3-down, 1-up tracking with a 2AFC task. With the transformed up-down, the number of reversals does not usually exceed the sixteen with at least four reversals run having a large step size or factor, and the remaining run having a small step size or factor. For a shorter transformed up-down tracking the user can opt for twelve reversals, four run with a large step size or factor and eight run with a small step size or factor. In both cases, the threshold should be calculated on the reversals run with the small step size or the small factor only.
Another point the user should pay attention to is the "distance" between the midpoint of the psychometric function underneath the experiment (e.g., 75% for 2AFC, 66% for 3AFC, 50% for yes/no task, etc.) and the point of the psychometric function that the staircase tracks (i.e.,

50% for method of limits and simple up-down, 71% for the 2-down, 1-up, 79% for the 3-down, 1-up, etc.). The staircase should be chosen in such a way that it tracks a point that is at least equal (or higher) than the midpoint of the psychometric function (e.g., Taylor, 1971; Green, 1990; Treutwein, 1995). In other words, if we are running a yes/no task (midpoint = 50%) any staircase can be used. On the contrary, if we are running a 2AFC (midpoint =75%) we should avoid both the method of limits, the simple up-down but also the 2-down, 1-up (Leek, 2001; Kollmeier et al., 1988). A final recommendation is to favour the comfort of the subject: the starting delta of the experiment should be sufficiently high for an easy first set of trials.

How to use the toolbox

The graphical interface of the toolbox is pretty straight forward. The fields the user has to fill in (or select) are labelled with the same words used in this text (e.g., standard, delta, factor, etc.). There are three levels uf users (see below) but the main difference when using the toolbox depends whether the user logs into the toolbox with a personal account or uses the toolbox like a guest.

Use of the toolbox with a personal account

The typical example of a user with a personal account is an experimenter or a teacher that needs to collect and save data from a group of participants. This user can register, create a personal account, then log in the toolbox. At this point, the user can use the toolbox as a guest (see below) or send an invitation code (or a link) to a third party user. This last user (e.g., the participant of an experiment) can paste the code into the interface of the toolbox and run the experiment. The data produced by this participant will be stored and available in the account of the registered user together with the personal data of the registered user.

Simple guest user

When the user connects to the main page, the first thing s/he has to do is to select the experiment s/he wants to run. When selecting the experiment, the user is then prompted with an interface that collects demographic details. The user needs to fill in at least the "Name" field (the only compulsory field) then s/he can move on and set the characteristics of the experiment. At the top, the user can set the characteristics of the standard tone, such as the intensity (expressed in dB Fs), the frequency and the duration of the standard tone (respectively in Hz and ms).

SuperUser

The superuser has the complete view of every account and all the data produced with the toolbox, both by registered users and guest users. The superuser can create other superussers.

Errors Index

Error 1
Error 2

…

Account privileges

In the following section will be explained which kind of users that are going to use this toolbox:

Guest

A guest user is rather someone who wants to run his own test and save his results on a local storage or someone who have been invited to run the test through an invite code (see the section invite code)

A guest can only run tests but can't save his settings and his personal information and doesn't have a history of his tests.

Account owner User

An account owner user (AoU) is someone who registered his own account. Every AoU has an invite code (or a link) that can be shared to every other user.
Every test done from the AoU or from users who used the AoU invite code are saved in his account test history, moreover every setting or personal information of AoU is saved.

Invite code

Guests and AoUs can use an invite code (or an invite link, they are equivalent) when running a test:

- If a Guest doesn't use it the results will be visible only on the results page right after the test is finished, once that page is closed nobody can see them (except for the SuperUsers);
- If a Guest uses an invite code the results will be visible on the results page and will appear on the test history of the invite code's AoU;
- If an AoU doesn't use an invite code and doesn't insert a name in the "demographic data" page, the test results will be visible after the test (in the results page) and in his test history (as done from him);
- If an AoU doesn't use an invite code but inserts a name, the test will be saved as done from a guest with the inserted personal informations who used the invite code of the AoU, so it will be visible in the test history of the AoU (as done from a Guest).
- If an AoU uses an invite code and doesn't insert a name, the test will be saved as done from a guest (with the AoU personal informations) who used the invite code, so it won't be visible in the test history of the AoU who did the test, but only in the one of the invite code's AoU;
- If an AoU uses an invite code and inserts a name, the test will be saved as done from a guest with the inserted personal informations who used the invite code, so it won't

be visible in the test history of the AoU who did the test, but only in the one of the invite code's AoU;

Datafiles returned by the toolbox

The toolbox returns two types of data files: a synthetic and a complete datafile. The reduced datafile includes the demographic details of the subject, and the threshold estimates for each block of trials. The complete datafile is a complete log of all the details of the experiment. For example, it includes the experiment name, the demographic detail of the subject, the characteristics of the standard tone and those of the threshold tracking, and the minute details of all events occurred in each trial of the experiment. Both files are cvs text files with the data organised in a square matrix by rows and columns.

References

Amitay, S., Irwin, A., Hawkey, D. J. Cowan, J. A., and Moore, D. R. (2006). “A comparison of adaptive procedures for rapid and reliable threshold assessment and training in naive listeners,” J. Acoust. Soc. Am. 119, 1616-1625.

Baker, R. J., & Rosen, S. (1998). “Minimizing the boredom by maximising likelihood – Efficient estimation of masked threshold.” Brit. J. Audiol., 32, 104-105.

Boring, E. G. (1961). “Fechner: inadvertent founder of psychophysics,” Psychometrika, 26, 3-8.

Brainard, D. H. (1997). “The Psychophysics Toolbox,” Spatial Vision, 10, 433-436.

Brown, L. G. (1996). “Additional rules for the transformed up–down method in psychophysics,” Percept. Psychophys. 58, 959-962.

Dixon, J. W., and Mood A. M. (1948). “A Method for Obtaining and Analyzing Sensitivity Data,” J. Am. Stat. Ass. 43, 109-126.

Faes, L., Nollo, G., Ravelli, F., Ricci, L., Vescovi, M., Turatto, M., Pavani, F. and Antolini, R. (2007). “Small-sample characterization of stochastic approximation staircases in forced-choice adaptive threshold estimation,” Percept. Psychophys. 69, 254-262.

Fechner, G. T. (1889). Elemente der Psychophysik (Breitkopf & Härtel, Leipzig).

Francart, T., van Wieringen, A. and Wouters, J. (2008). “APEX 3: a multi-purpose test platform for auditory psychophysical experiments,” J. Neurosci. Meth. 172, 283-293.

García-Pérez, M. A. (1998). “Forced-choice staircases with fixed step size: Asymptotic and small-sample properties,” Vis. Res. 38, 1861-1881.

García-Pérez, M. A. (2002). “Properties of some variants of adaptive staircases with fixed step sizes,” Spatial Vision, 15, 303-321.

García-Pérez, M. A. (2009). “Denoising forced-choice detection data,” Brit. J. Math. Stat. Psy. 73, 75-100.

Gescheider, G. A. (2003). Psychophysics: the fundamentals (Lawrence Erlbaum Associates, Hillsdale).

Geurts, L., Vouters, J. (2000). “A concept for a research tool for experiments with cochlear implant users,” J. Acoust. Soc. Am. 108, 2949-2956.

Grassi, M., Soranzo, A. (2009). “MLP: a MATLAB toolbox for rapid and reliable auditory threshold estimations,” Behav. Res. Meth. 41, 20-28

Green, D. M. (1990). “Stimulus selection in adaptive psychophysical procedures,” J. Acoust. Soc. Am. 87, 2662-2674.
Green, D. M. (1993). “A maximum-likelihood method for estimating thresholds in a yes-no task,” J. Acoust. Soc. Am. 93, 2096-2105.
Green, D. M., & Swets, J. A. (1966). Signal detection theory and psychophysics. New York: Wiley.
Kollmeier, B., Gilkey, R. H., and Sieben, U. K. (1988). “Adaptive staircase techniques in psychoacoustics: A comparison of human data and mathematical model,” J. Acoust. Soc. Am. 83, 1852-1862.
Leek, M. R. (2001). “Adaptive procedures in psychophysical research,” Percept. Psychophys. 63, 1279-1292.
Levitt, H. (1971). “Transformed up–down methods in psychoacoustics,” J. Acoust. Soc. Am. 49, 467-477.
Marvit, P., Florentine, M., and Buus, S. (2003). “A comparison of psychophysical procedures for level-discrimination thresholds,” J. Acoust. Soc. Am. 113, 3348-3360.
Pelli, D. G. (1997). “The VideoToolbox software for visual psychophysics: Transforming numbers into movies,” Spatial Vision, 10, 437-442.
Pentland, A. (1980). “Maximum-likelihood estimation: The best PEST,” Percept. Psychophys. 28, 377-379.
Saberi, K., & Green, D. M. (1997). “Evaluation of maximum-likelihood estimators in nonintensive auditory psychophysics.” Percept. & Psychophys., 59, 867-876.
Schlauch, R. S., and Rose, R. M. (1990). “Two-, three-, and four-interval forced-choice staircase procedures: Estimator bias and efficiency,” J. Acoust. Soc. Am. 88, 732-740.
Soranzo, A., and Grassi, M. (2014). “PSYCHOACOUSTICS: a comprehensive MATLAB toolbox for auditory testing”. Front. Psychol., 5, 712.
Stanislaw, H., & Todorov, N. (1999). “Calculation of signal detection theory measures.” Behav. Res. Meth., Instr., & Comp., 31, 137-149.
Taylor, M. M. (1971). “On the efficiency of psychophysical measurement,” J. Acoust. Soc. Am. 49, 505–508.
Taylor, M. M., and Creelman, C. D. (1967). “PEST: Efficient estimates on probability functions,” J. Acoust. Soc. Am. 41, 782-787.
Treutwein, B. (1995). “Adaptive psychophysical procedures,” Vis. Res. 35, 2503-2522.
von Bekesy, G. (1947). “A new audiometer,” Acta Otolaryngology, 35, 411-422.
Watson, A. B., and Fitzhugh, A. (1990). “The method of constant stimuli is inefficient,” Percept. Psychophys. 47, 87-91.
Watson, A. B., and Pelli, D. G. (1983). “QUEST: A Bayesian adaptive psychometric method,” Percept. Psychophys. 33, 113-120.
Wetherill, G. B., and Levitt, H. (1965). “Sequential estimation of points on a psychometric function,” Brit. J. Math. Stat. Psy. 18, 1-10.
Wichmann, F. A., and Hill, N. J. (2001). “The psychometric function: I. Fitting, sampling, and goodness of fit.” Percept. Psychophys., 63, 1293-1313.