Tudor Popescu
MRes Thesis

The plasticity of the human perceptual systems in processing auditory- and spatial-frequencies

Supervisor: Dr Hannah Smithson

Durham University
Department of Psychology

30th March, 2009

Table of Contents

1. INTRODUCTION

1.1. Overview

Perceptual processing occurs in a hierarchical fashion: for example in vision, light enters the eye and is processed by photoreceptor cells; the neural signal is then transmitted through layers of the retina, before it is sent to the Lateral Geniculate Nucleus (LGN) and onto the layers of the visual cortex. Similarly, in hearing, sound waves enter the outer ear and reach the eardrum, which transmits information to the ossicles of the middle ear, and then to the basilar membrane of the cochlea; finally, the sound information is converted by hair cells into electrical nerve impulses which travel to the auditory cortex.

One interesting metaphor for the human perceptual processing is to consider it as an artificial information processing system, constructed to extract useful information from the input. In engineering terms, we might liken this to a series of filters, i.e. devices that are *selective* (in the sense that they respond only to a *subset* of all possible inputs), and are thus able to extract relevant properties of the input. Such filters have been shown to exist in *vision* and *hearing*, selectively processing the input *image* or *sound* in terms of its frequency components.

There is, for example, physiological evidence that there are nerve fibers that respond to a subset of auditory frequencies in the cochlea. Using the above-mentioned metaphor, each of these nerve fibers is conceptualised as a filter that functions rather like the different channels of equalisers found in stereo systems.

However, this view of our perceptual systems leaves no room for flexibility; the frequency-selective filters are conceptualised as somewhat rigid, hard-wired units at a low level in our sensory systems.

An important question to ask in psychology is whether we can get better at performing a particular perceptual task with extended practice, and there is a lot of evidence that *perceptual learning* does take place. This has been confirmed in the case of extended practice with simple stimuli. What I am essentially asking in this thesis is whether we can also learn about more abstract stimuli such as noise, and I intend to answer this question by performing two related experiments, one in vision and one in hearing.

1.2. Structure of thesis

In the material to follow, I will start by shortly reviewing, in Chapter 2, the methods that are commonly used in *psychophysics*, with an emphasis on the particular methods that I have used in my experiments. The notions defined in this chapter are needed for the understanding of the following chapters.

Moving on to Chapter 3, I will start establishing the connection between perceptual processing and the above-mentioned engineering-like structures, by introducing the notion of *frequency-selective channels*, which are believed to exist in both *vision* and *hearing*. I will then review the evidence that suggests these structures are not *completely* rigid by discussing *perceptual learning*, in Chapter 4.

Chapter 5 will put the literature review done by the previous two chapters into perspective, and will explain what my motivation is for carrying out the present study. Chapter 6 will then give a description of the stimuli that I have employed, which are mathematically very similar for the two experiments.

Further on in Chapter 7, I will describe the design and methodology for the vision experiment, and I will present its results both by directly plotting the data *and* by doing multiple statistical analyses that examine trends in the data. Those analyses will be able to validate or invalidate my initial hypotheses, as well as reveal new results that were not foreseen. Chapter 8 does the same for the hearing experiment. Each of those chapters ends with a discussion.

Finally, I conclude with Chapter 9, that discusses the overall conclusions that can be drawn from the entirety of this work, as well as suggest similar research questions that might be later pursued.

2. BACKGROUND TO PSYCHOPHYSICAL METHODS

2.1. Signal detection theory

Signal detection theory (SDT) is a mathematical technique developed in the 1950s and concerned with measuring the way decisions are taken – by humans or machines – in the process of detecting a target signal under conditions of uncertainty, for example in the presence of noise.

SDT was first developed and used in radio communications but has in the meantime found applications in several fields where it is necessary to discern between a "useful" signal (the *target signal* or *target*) and a "parasite" signal (*noise*).

In psychophysics, the assumption made by SDT is that the human observer performing a task under uncertainty conditions makes judgements and decisions based on his perceptions; performance on a task can be related to the underlying neural mechanisms with the aid of the mathematical model proposed by SDT.

If we consider the case of a detection task, in which an observer is required to detect the presence of a target embedded in noise[1], then the two intervals of the trial, the "noise alone" (N) interval and the "target+noise" (T+N) interval, can be represented by their probability distributions along a horizontal axis – called the *perceptual decision axis* – which represents a combined measure of the signal's characteristics, including contrast. Both the target and the noise are projected onto this axis, and their distributions are usually both considered to be normal.

The observer makes decisions regarding the presence or absence of the target based on a simple rule – if the measure of the signal along the perceptual decision axis is greater than a certain criterion level, then the observer will respond that the target signal *was* present. The position of the criterion along the axis depends on several factors, including the observer's competence at the task and the design of the experiment.

[1] The noise can be either internal to the observer (e.g. spontaneous neural discharge) or external, i.e. added to each interval by the experimenter.

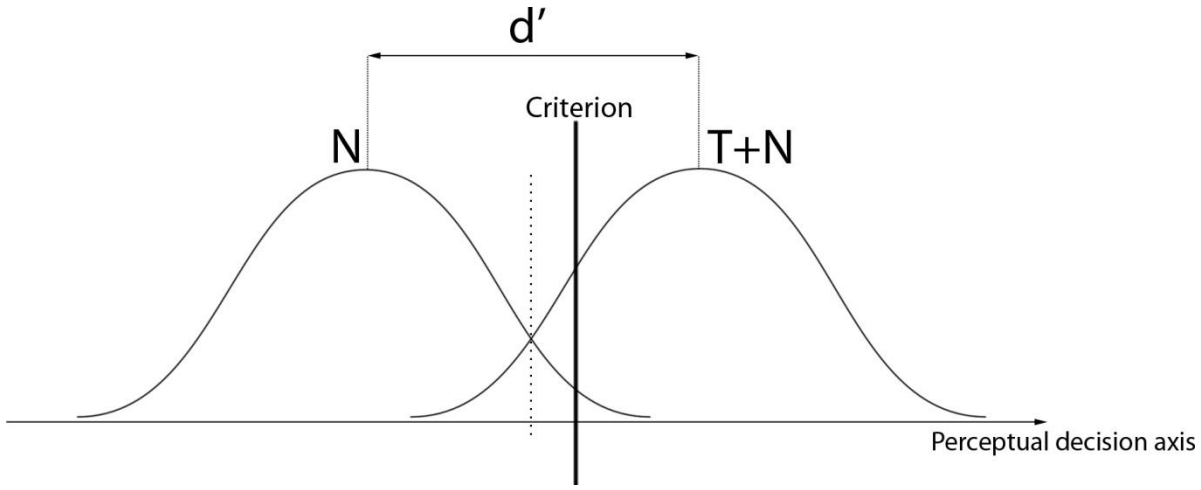

Figure 1: The classic representation of the principles of signal detection theory. The distribution corresponding to the target+noise interval (T+N) is to the right of that corresponding to the noise interval (N), because its overall contrast is higher than that of the noise alone. The criterion represented by the continuous vertical line is conservatively-biased because it stands to the right of the dotted line that crosses through the point where the two probability distributions meet, and that represents the *neutral criterion.*

Based on the response given by an observer in a detection task (absence or presence of target) and on reality (whether or not the target was actually present), trials can be sorted into one of four categories:

Table 1: The four possible trial categories, based on whether or not the target was present and on whether or not the target was reported as being present.

	Response: "target absent"	*Response: "target present"*
Reality: target absent	Correct rejection	False alarm
Reality: target present	Miss	Hit

Based on the proportion of each of those four types of trials throughout an experiment, the following statistics can be derived:

- the *sensitivity index*, **d'** – gives a numerical measure of the observer's *sensitivity*; d' is the minimum distance that needs to exist between the N distribution and the T+N distribution in order for the target to be detected correctly by the observer;
- the *likelihood ratio*, **β** – gives a numerical measure of the observer's a priori *bias* towards one response or the other. The observer's *bias* is independent of his *sensitivity*.

An observer's bias can be influenced by a system of punishments and rewards associated with the experiment, which sets how important it is that the observer does not *miss* and, on the other hand, how important is it that he does not give *false alarms*. The resulting bias can be described, in the two extreme cases, as follows:

- a *liberal* bias, when the cost of a *miss* is high, for example when a policeman has to decide whether or not to stop and search a suspect for weapons; the observer would rather give a *false alarm* rather than *miss*

- a *conservative* bias, when the cost of a *false alarm* is high, for example the decision of "crying wolf" too often, which eventually leads to no one coming to the rescue; the observer tries to avoid *false alarms*, even at the risk of *missing* a few targets (i.e. having to solve the problem on his own)

A *neutral criterion* (i.e. no bias) is one that does not favour any one option a priori, and can be schematically represented by the vertical line crossing through the intersection point of the two distributions. There are methods that encourage observers to set a neutral criterion; among those is the two-alternative forced-choice (2AFC), which I will be using in my experiments and which is defined in section 2.2.2.1.

In each of the two intervals[2] of a trial, the observer either sees noise alone or target+noise, and has to choose whether the target was present in the interval he just saw. In an interval where the target was present (a situation described by the T+N curve in Figure 2), the area under the T+N curve which is to the right of the criterion level represents the *hit rate*, because for that part of the curve, the observer's response was correct, i.e. that there was a target present. The area under the curve that is to the left of the criterion represents the *miss rate*, because there the observer failed to identify the existing target.

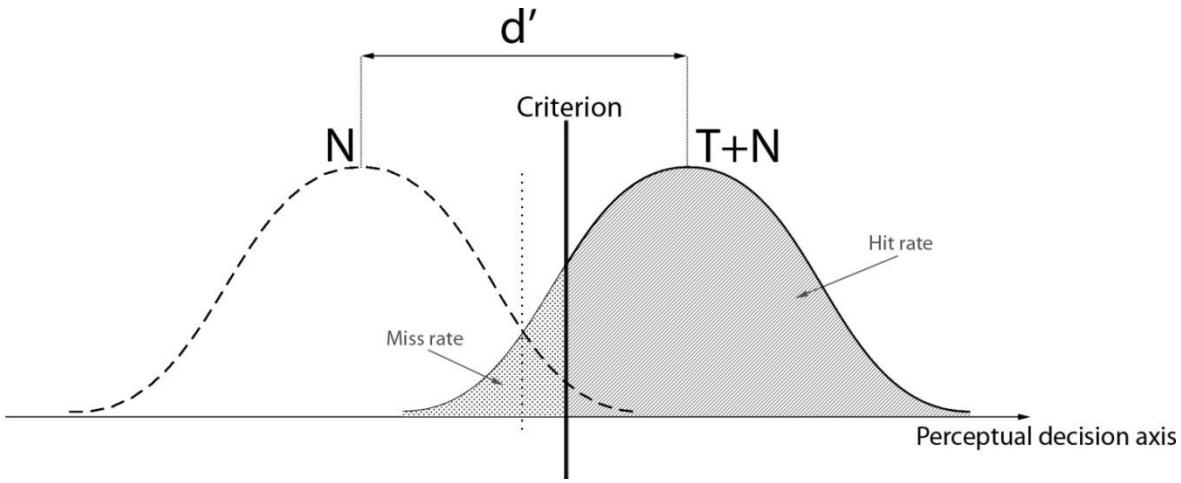

Figure 2: Graphical description of an interval that contains target+noise, the distribution of which is represented with a continuous line along the perceptual decision axis. The distribution of a noise alone interval is also included on the graph, for reference, and plotted using a dashed line. The dotted vertical line that crosses through the intersection of the two curves represents the neutral criterion (no bias). The areas under the T+N curve that quantitatively represent the miss and hit rates are emphasised with different patterns.

Similarly, the *correct rejection* and *false alarm* rates can be defined graphically for intervals where the target was not present, as seen Figure 3 below.

[2] There are two intervals in a 2AFC-type trial (see section 2.2.2.1 on 2AFC), however in the most general case, a trial can have any number of intervals.

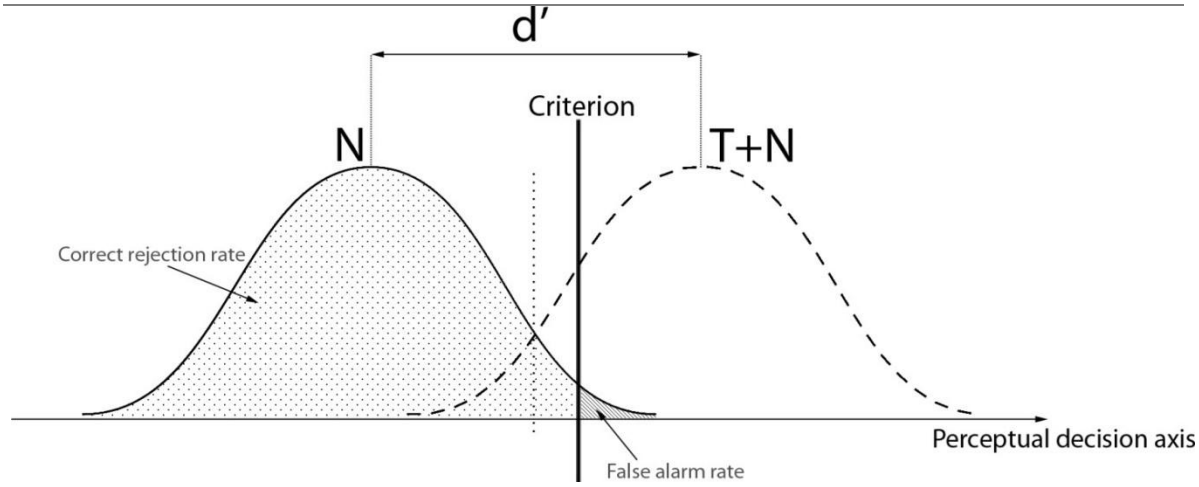

Figure 3: Graphical description of an interval that contains noise alone (no target), the distribution of which is represented with a continuous line along the perceptual decision axis. The distribution of a target+noise interval is also included on the graph, for reference, and plotted using a dashed line. The dotted vertical line that crosses through the intersection of the two curves represents the neutral criterion (no bias). The areas under the N curve that quantitatively represent the correct rejection and false alarm rates are emphasised with different patterns.

I will make use of the notions defined by SDT later, when I define my research hypothesis.

2.2. Experimental procedures used in psychophysics

2.2.1 Methods of measurement and psychometric functions

The *detection threshold* (or simply *threshold*) of a signal is the critical value of the measure of its strength above which the signal can be detected by a particular observer and below which it cannot. For example, for a visual sinusoidal grating, whose *contrast* is a measure of its strength, its detection threshold (or: contrast threshold) is the contrast value for which the signal is just starting to become distinguishable.

According to SDT, the threshold value is determined by the observer's criterion and the separation of the N and T+N distributions. Therefore, on a particular trial (which is just one sample from the relevant distribution), threshold measurement has a random character and does not reflect the strength value above which the signal is *always* detected. Instead, one has to "define" threshold in terms of the performance level at which one desires to measure it, and decide upon a *percentage of cases* in which correct identification is achieved (60%, 75% and 90% are typical values). This percentage is then used in the measuring procedure, which produces a threshold value; for example, if using the *method of constant stimuli*, the strength of the signal will be held constant for several trials, and the percentage of correct responses will be calculated from those trials at a fixed stimulus level. The operation is then repeated for several values of signal strength, thus determining thresholds for several performance levels (i.e. percentage of correct responses). Finally, a curve is fitted to those points, using one of several possible models. This curve is called a psychometric function and describes the performance of an observer in a particular sensory task, at each possible performance level. The psychometric function allows one to read the threshold value for any performance level, for example,

60%, 75% and 90% correct responses. Figure 4 below gives an example of how a psychometric function might look.

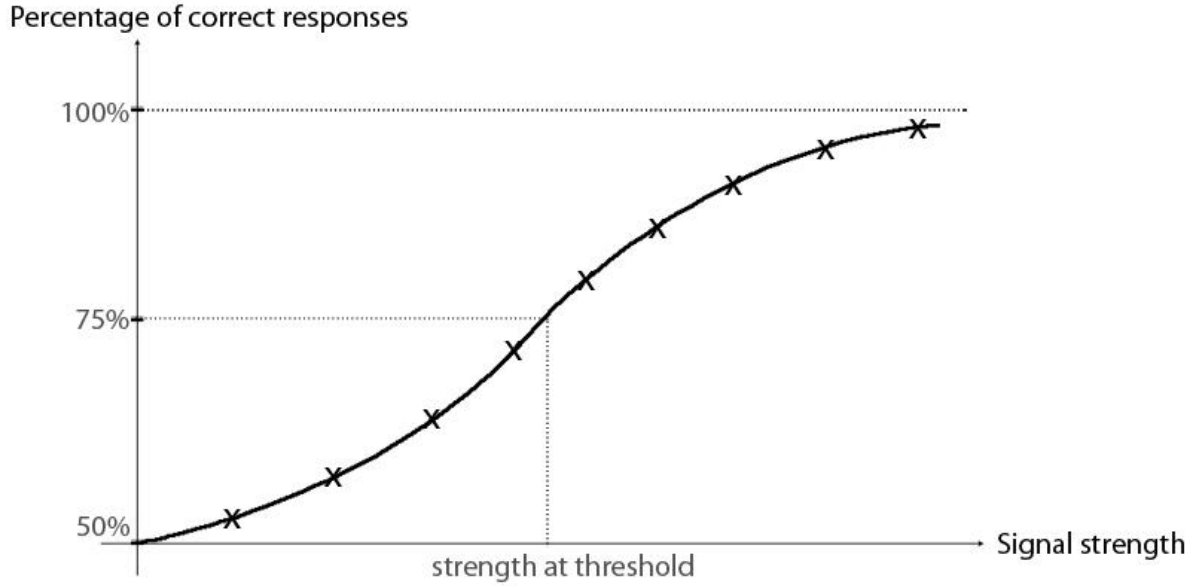

Figure 4: A typical psychometric function shape, obtained by fitting a curve through several data points (marked by crosses) that were obtained for different "percentage correct" performance levels. The psychometric function allows one to read the threshold value for any performance level (e.g. 60%, 75% or 90%); in this qualitative representation, I have chosen "threshold" to mean "responses correct 75% of the time". The point on the curve corresponding to the threshold may or may not be one of the data points that the curve originates from. The y-axis' intersection with the x-axis occurs at 50% correct responses because even in the absence of a signal, the percentage of correct responses cannot drop below the 50% corresponding to pure chance.

An alternative to the method of constant stimuli is a set of procedures known as *adaptive methods* or *staircases*; in such a procedure, rather than choosing a fixed set of contrast levels at the start of the experiment, contrast for a particular trial is determined by the observer's previous responses. Typically, contrast is decreased after correct responses and increased after incorrect responses. The rule for adjusting contrast in the staircase is chosen such that the staircase converges on a particular predefined performance level.

Threshold is normally defined in the context of *detecting* a signal (i.e. the non-signal interval is a uniform field), but a threshold can also be defined when the correct identification of the signal involves distinguishing it from another signal, i.e. as part of a *discrimination* task. The necessary threshold in this case is greater than in the detection case, because, according to Weber's law, as the background increases in strength, so does the threshold. Section 3.4.3 will discuss in more detail what exactly makes a task "cross the line" from *detection* to *discrimination*.

2.2.2 Methods of eliciting responses from observers

2.2.2.1 Two-alternative forced-choice (2AFC)

The simplest way of eliciting responses from an observer is to present only one stimulus per trial (as part of a single interval), and, after each trial, to ask the observer whether they saw the stimulus or not. The observer is thus required to reply either "yes –

the stimulus was present", or "no – the stimulus was not present". However, by using this method, we allow the observers' bias to influence the results; for example, some observers might have a tendency to give affirmative answers more often than others.

For this reason, a *forced-choice* method is often preferred instead. This involves having *several* stimulus-containing intervals as part of each trial, rather than only one interval, as in the "yes/no" method. Only one of these stimuli contains the target, and having been assured of this fact, the observer is required to determine which interval (stimulus) contained the target. A special case of the forced-choice method is the *two-alternative* forced-choice (2AFC), whereby the observer is asked in which of the *two* stimuli (intervals) he just saw lies the target..

The main difference between using the "yes/no" method and a forced-choice method is that in the latter, the observer knows for sure that the target was in one of the intervals presented during the trial. On the other hand, in the "yes/no" method, the observer must decide whether or not a target is presented, and this decision can be influenced by two factors: the perceptual strength of the stimulus and the observer's willingness to say "yes". The latter is what constitutes the bias, which the forced-choice method helps eliminate.

As part of a forced-choice method, the different stimuli can be either *spatially* or *temporally* separated; that is, they can be all simultaneously visible in different parts of the screen (*spatial* AFC task) or they can be presented one after the other throughout the length of the trial (*temporal* AFC). The choice of either a spatial or temporal AFC is usually determined by the nature of the experiment or by the limitations of the equipment. Sometimes, both alternatives might be suitable, although care should still be taken over which one should be chosen, as they have been reported to lead to potentially different results, under comparable conditions (Peli, Garcia-Perez, Giorgi, & Woods, 2004).

For my experiments, I will be using a *temporal 2AFC* method.

3. FREQUENCY-SELECTIVE CHANNELS IN VISION AND HEARING

3.1. Spatial-frequency

It seems natural to describe images in terms of various properties which can be easily observed across space, such as hue, saturation and luminance. It seems less intuitive though, at least at first glance, that we could also consider them in terms of *frequencies*. However, each of these "easily observed" properties can be seen as either functions of *space* or functions of *spatial-frequency*.

Spatial-frequency is defined as the rate at which some property – normally *luminance* – varies across a given area of space. Thus, in an image of a given size, more cycles of *luminance variation* means higher spatial-frequency. This is exemplified in Figure 5 with two simple *sinusoidal gratings*, i.e. images whose luminance varies sinusoidally along one direction (in this case: the *vertical* direction).

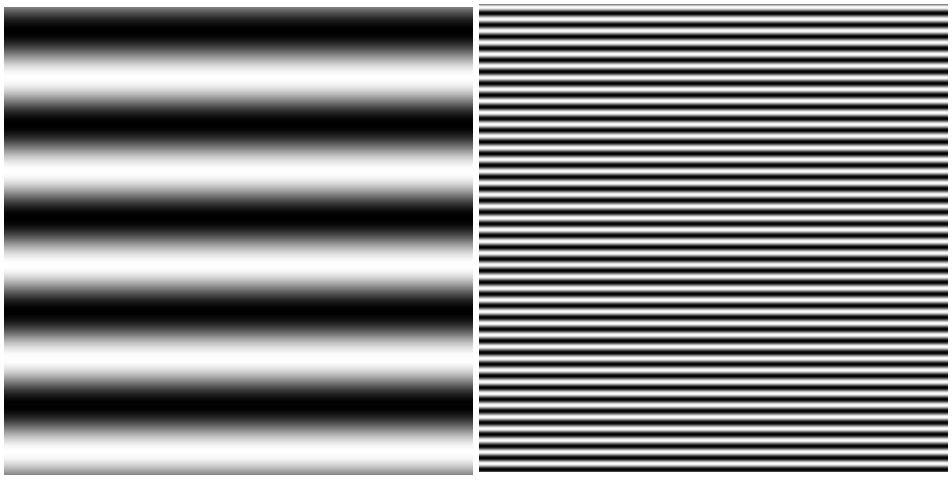

Figure 5: Left – an image with few bars of luminance (low spatial-frequency). Right – an image of the same size with many bars of luminance (high spatial-frequency).

Spatial-frequency is usually measured in *cycles per degree of visual angle* (*cpd*), or in *cycles per image* (*cpi*). The word "cycle" refers to a complete sinusoidal cycle that the luminance performs in the image. Also a measure of spatial-frequency is the *visual field*

angle (or: the *angular subtense*) that one cycle of grating subtends, as measured at the observer's eye. Technically, the *angular subtense* is the reciprocal of *spatial-frequency*, similar to how the *period* is the reciprocal of *frequency*.

The notion of spatial-frequency is normally used in psychology by assuming that the visual system is working as an approximately linear system, which means that if we know the response to one spatial-frequency component, we can predict the response to other spatial-frequency components.

3.1.1 Fourier decomposition in terms of spatial-frequency

There is a mathematical relationship between a time series (different values at different points in time) and the frequency domain (different values for different frequency components). If s(t) is the time series, then its Fourier transform, F(f), a function of frequency, is defined as:

$$F(f) = \int_{-\infty}^{\infty} s(t)\, e^{-2\pi j f t}\, dt$$

The function of time s(t) can be then reconstructed from the Fourier transform[3]:

$$s(t) = \int_{-\infty}^{\infty} F(f) e^{2\pi j f t}\, df$$

The most common examples of Fourier analysis relate signals that vary in *time* – of which *sounds* are the most representative – to their *temporal-frequency* (or: *auditory frequency*) representation, with *time* being measured in *seconds* and *temporal-frequency* being measured in *cycles per second*. This is the case described at the beginning of this section. However, with the same mathematical operations, signals that vary in *space* – such as images[4] – can also be related to their *spatial-frequency* representation. Pursuing the analogy, space can be measured in *degrees of visual angle*, and spatial-frequency is measured in *cycles per degree of visual angle*.

The *spatial-frequency* representation of an *image* (or the *temporal-frequency* representation of a *sound*) is an equivalent of the *spatial* representation of the *image* (or the *time*-varying representation of the *sound*); no information is lost by transitioning between the two domains in either direction; we are simply representing the same information in a different way.

As far as the Fourier analysis is concerned, a sine-wave is the simplest possible signal (visual or auditory), because it only has one frequency component. Fourier analysis allows any complex wave to be decomposed in a sum of sines and cosines, in much the same way that a musical chord can be described of consisting of the several notes that are played simultaneously as part of that chord.

[3] This relation defines the *inverse Fourier transform*.

[4] A visual signal can, however, also be a function of space, if we follow its evolution in time. I will not be concerned with this aspect in my thesis.

The result of Fourier decomposition is a set of *sinusoids* (or: *sine-waves*) of varying amplitudes and phases, over a range of frequencies whose limits are set by the Nyquist-Shannon theorem, which states that an analogue signal (i.e. one that is continuous in time) that has been sampled can be perfectly reconstructed from the samples if the sampling rate exceeds $2F_{max}$ samples per second, where F_{max} is the highest frequency in the original signal[5].

A consequence of the Nyquist-Shannon theorem is that a discrete image consisting of an NxN matrix of pixels can be represented – without any loss of information – in the frequency domain, with discrete spatial-frequency values ranging from $-\frac{N}{2}$ to $\frac{N}{2}$. When doing the Fourier analysis of an image, our primary concern is to look at the energy distribution of the various spatial-frequencies. This distribution is derived from the amplitude spectrum – the plot of the *absolute values* of the complex numbers[6] returned by the Fourier transform. A plot of the *phases* of those complex numbers would constitute the *phase spectrum*, which, in the context of my area of study, is not particularly relevant.

3.1.2 Spatial-frequency in natural images

One can identify regions of low or high spatial-frequency in any type of image, whether artificially created (such as Figure 5) or natural. For instance, in a scene with pebbles on a beach (Figure 6), if we look to the pebbles which are closer to us, the fact that we see fewer of them and that they appear bigger means that the image has a low spatial-frequency, whereas if we look further to pebbles which are more distant, we can see more pebbles which appear smaller, i.e. higher spatial-frequency.

Figure 6: Pebbles on a beach[7].

[5] Nyquist–Shannon sampling theorem. (2008, October 28). In Wikipedia, The Free Encyclopedia. Retrieved 19:02, October 29, 2008, from http://en.wikipedia.org/w/index.php?title=Nyquist%E2%80%93Shannon_sampling_theorem&oldid=247925369

[6] A complex number is one that can be written in the form **a** + **bi**, where **i** is called the *imaginary unit* (defined by $\mathbf{i}^2 = -1$) and where **a** and **b** are real numbers called the *real part* and the *imaginary part* of the complex number, respectively.

[7] Retrieved April 4, 2008, from http://www.seafriends.org.nz/oceano/f210431t.jpg

The degree to which low and high spatial-frequencies are present in a natural image tells us something about that image (Bar, 2004):

- *high* spatial-frequencies indicate that there are abrupt luminance variations (e.g. edges) in the image, and/or fine details, such as sand on a beach;
- *low* spatial-frequencies indicate smooth transitions (e.g. blurred edges) and the lack of fine details.

Analogies between natural and artificial images in the space domain may sometimes lead to valid conclusions about how the Fourier decomposition is going to look. For instance, a scene with sand on a beach is very similar to how a "noisy" image looks, and indeed the Fourier equivalent of such a scene is a relatively flat spectral distribution, as is the case with an image representing *white noise*[8].

The Fourier transform of an image representing a natural scenes is often a relatively linear decreasing curve, when plotted a double-logarithmic scale (Tolhurst, Tadmor, & Chao, 1992). This shows that there is more energy in the image at the lower spatial-frequencies, rather than at the higher spatial-frequencies.

3.2. The contrast sensitivity function and the channels model

One of the factors that determine the visibility of a grating is its contrast[9], and the contrast required for detecting a grating (the *contrast threshold*[10]) depends on its spatial-frequency. This dependence is usually presented in the form of a plot of *sensitivity*[11] as a function of *spatial-frequency*, and this curve is called the contrast sensitivity function (CSF), a qualitative plot of which is shown in Figure 7.

[8] White noise is noise that has a flat spectral disctribution.

[9] I use the term "contrast" in the Michelson sense, i.e. $\frac{I_{max}-I_{min}}{I_{max}-I_{min}}$, where I_{max} and I_{min} represent the highest and lowest luminance respectively.

[10] The contrast threshold is the minimum value of contrast that an image needs to have in order to be detectable.

[11] Here the word "sensitivity" refers to contrast sensitivity, which, in turn, is the reciprocal of contrast threshold. In other words, "Contrast sensitivity is the reciprocal of the contrast of a grating at the threshold of visibility." (Stromeyer III & Julesz, 1972)

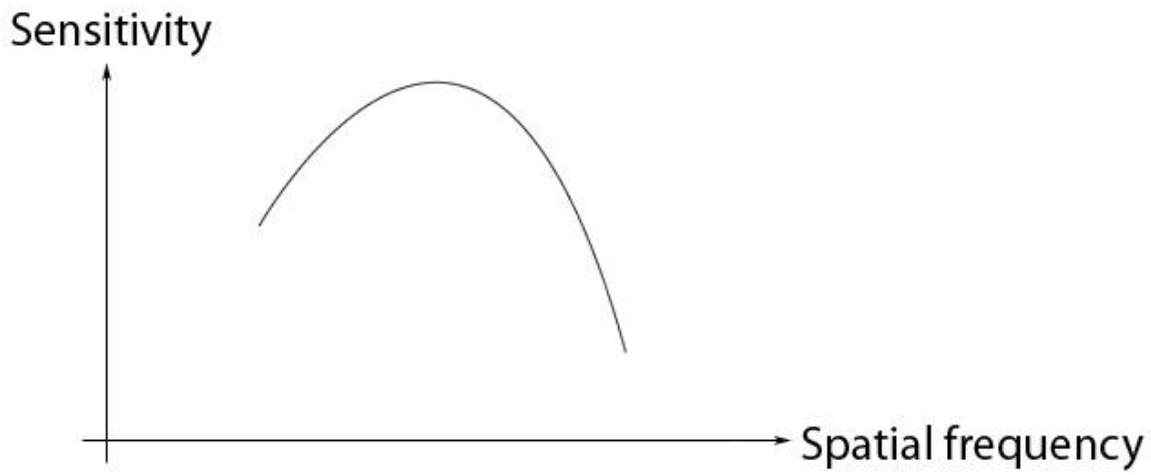

Figure 7: Qualitative representation of the contrast sensitivity function (CSF), which shows how sensitive we are to the various spatial-frequencies. Note that there are spatial-frequencies which are outside the range of the CSF and therefore those spatial-frequency components cannot be seen by the human eye.

The first research that looked into the contrast sensitivity function was that of Schade (1956), who regarded the visual system as an optical system and expressed its contrast sensitivity as a function of the photoelectric transfer characteristics of the receptors of the visual system (the rods and the cones).

As reviewed in (De Valois & De Valois, 1980), early studies in spatial vision considered the contrast sensitivity function to be a single channel. This view was challenged for the first time in (F. W. Campbell & Robson, 1968), where it has been suggested that the human CSF represents the envelope of sensitivity of *many* underlying channels, each of which is narrowly tuned to a small range of spatial-frequencies, the range being centred around one "optimum" spatial-frequency, to which the channel is most sensitive to. It could be put (F. W. Campbell & Robson, 1968) that each channel has its own contrast-sensitivity function, and the envelope to all of them is the "global" contrast-sensitivity function, which describes the visual system as a whole. The channels on the other hand show the sensitivity of each subrange of spatial-frequencies and are called *selective channels*, or *selective filters*, in the sense that they are only sensitive to *one* range of spatial-frequencies and insensitive to all others.

The sensitivity curves of these underlying channels are represented in Figure 8 below, together with the overarching curve of the CSF.

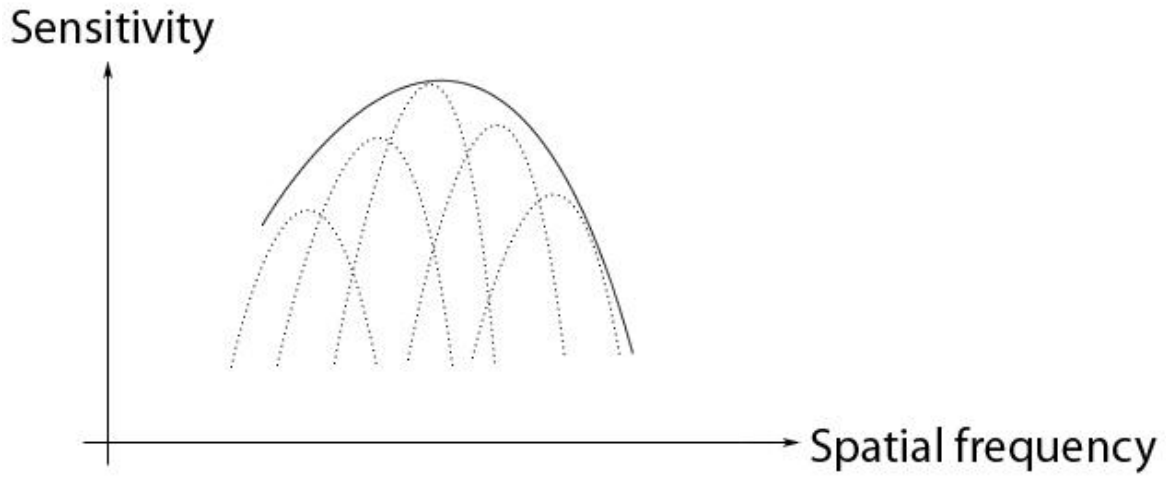

Figure 8: Qualitative representation of the CSF (continuous line) and the underlying spatial-frequency-selective channels (dotted lines).

3.2.1 Evidence for the existence of spatial-frequency-selective channels in vision

In the following sections I will review the classical evidence that supports the existence of spatial-frequency-selective channels in the visual system. I will consider four main lines of evidence (Williams, 2004).

3.2.1.1 Detection thresholds for complex wave forms

The CSF provides a concise description of the relative visibility of patterns whose luminances vary sinusoidally at different spatial-frequencies.

Campbell and Robson (1968) posed the question of whether the visibility of complex wave forms can be predicted by knowing the visibility of the individual sinusoidal components (as calculated by Fourier decomposition). This type of question sits within the framework of *linear systems analysis*.

A *linear system* is one that can be described mathematically using a linear function of the form

$$f(x) = ax+b,$$

where x is the input and f(x) is the output. A linear system is one that satisfies the property of *superposition*, meaning that with several signals at the input, the output of the system is the sum of the responses that would be produced by each input signal individually. To express this formally, assume that $x_1(t)$ and $x_2(t)$ are the input signals, and that $y_1(t)$ and $y_2(t)$ are the system's responses to them:

$$y_1(t) = H \{ x_1(t) \}$$

$$y_2(t) = H \{ x_2(t) \},$$

then the superposition of x_1 and x_2 (possibly scaled by factors α and β respectively) at the input of the system should produce the same response as the scaled sum of the individual responses y_1 and y_2:

$$\alpha y_1(t) + \beta y_2(t) = H \{ \alpha x_1(t) + \beta x_2(t) \}.$$

The critical point here is that if the visual system can be approximated as a linear system, then the superposition principle might allows us to predict the response to complex wave forms from knowledge of the response to simple sinusoids.

Campbell and Robson (1968) tested this empirically: by using the method of adjustment, they measured contrast thresholds for various luminance patterns. One critical comparison was between thresholds for simple (sine) gratings and for square-wave gratings[12]. The theory of Fourier decomposition tells us that the coefficients of the fundamental frequencies of these two types of gratings differ by a factor of $\frac{4}{\pi}$; indeed, the contrast threshold for the square-wave was found to be $\frac{4}{\pi}$ times smaller than that of the sine.

For a large range of spatial-frequencies, the energy contained in the higher harmonics of the square-wave grating (i.e. 3f and higher) did not contribute to detection (that is, did not lower the threshold), which implied that those harmonics were stimulating different underlying mechanisms. At spatial-frequencies which were sufficiently low, the higher harmonics of the square-wave started to contribute to detection, and, as a result, sensitivity to the square-wave increased.

What this seemed to suggests was that the visual system analyses images not only based on spatial cues such as edges but also by means of Fourier analysis done on the image, thus breaking it down into different spatial-frequency components which are then processed individually by the appropriate channels.

In addition to looking at threshold phenomena, Campbell and Robson (1968) also considered discrimination at suprathreshold contrasts; in particular, they were interested in predicting when a complex wave form would become visibly different from a sinusoidal wave form. Again, linear theory makes a specific prediction here: the square should become distinguishable from the sine grating when its third harmonic reaches its own threshold; if the system were not linear, the fundamental and harmonic terms would not simply add. Their results were in line with the linear systems theory which has been assumed so far.

3.2.1.2 Spatial-frequency adaptation

Using single-cell recordings, Campbell, Cooper and Enroth-Cugell (1969) found that *cat* neurones respond selectively to narrow ranges of spatial-frequency. In the same year, Blakemore and Campbell (1969) have studied – using psychophysical measurements – the properties of such spatial-frequency-selective neurones in *man*, and compared them with those found by Campbell et al. (1969) in the *cat*.

Blakemore and Campbell (1969) discovered the existence of a spatial adaptation effect, whereby if an observer became adapted (as a result of prolonged viewing) to a certain spatial-frequency, this caused a decrease in sensitivity only over a limited range of neighbouring frequencies centred around that adapting frequency, rather than over the whole range of the CSF. Observations of the occipital evoked potentials after adaptation suggested that the adaptation phenomenon was associated with a decrease in neural

[12] A square-wave grating is one whose luminance varies according to a hard-edged distribution rather than to a smooth, sinusoidal distribution, as in the case of sine gratings.

activity. This was accompanied by a subjective elevation of threshold (i.e. fading) for the test grating.

The said adaptation effect manifested two important properties:

- *orientational specificity*: a threshold elevation was only produced if the viewed stimulus had the same orientation as the stimulus used for adaptation; and
- *binocular transfer*: if only one eye was adapted, the other eye also caused a rise in threshold, if to a lower extent than that of the adapted eye.

These properties led Blakemore and Campbell (1969, p. 257) to believe that the adaptation effect "exposes the properties of central human neurones beyond the optic radiation", and thus that the locus where this effect takes place is *central* to the visual system.

No after-image of the grating was perceived by the observer when a uniform surface was presented after adaptation, which elliminated a conventional after-image phenomenon as the cause of the adaptation. Furthermore, the adaptation existed even when the observer did not fixate steadily upon the adapting (high contrast) grating.

As a result of the adaptation phenomenon, the CSF curve did not lower uniformly, but only in a narrow vicinity around the adaptation frequency, as can be seen in the qualitative graph displayed in Figure 9. The adaptation was subsequently done for various frequencies, and the depressed sensitivity range always centred around the adapting frequency. This suggests that only the sensitivity of a structure which is more narrowly tuned than the CSF (i.e. a *spatial-frequency-selective channel*) is affected by adaptation.

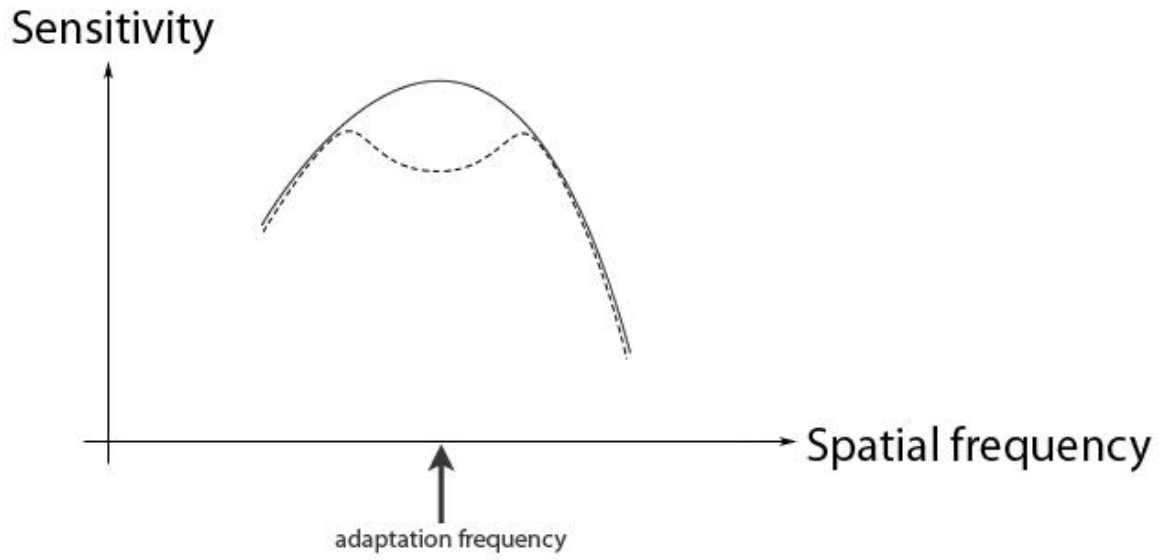

Figure 9: CSF *before* adaptation (continuous line) and *after* adaptation (dashed line); graph loosely adapted from (Blakemore & Campbell, 1969).

The elevation of threshold was deemed to be dependent upon the contrast of the adapting grating, whatever its spatial-frequency. The spatial-frequency selectivity of the adapted mechanism could be inferred by comparing the CSF *before* adaptation with the CSF *after* adaptation. However, the shape of the latter was found to be independent of both the spatial-frequency of the adapting grating *and* its contrast.

The adaptation effect was also emphasised in the context of complex gratings, which were analysed in terms of their Fourier decomposition, as suggested in (F. W.

Campbell & Robson, 1968). Both theories converged when, as a result of adaptation to a square-wave grating, there was a significant elevation in threshold for a sine-wave that had the same frequency as the third harmonic[13] of the adapting square-wave. This showed how adaptation to a *single* grating can simultaneously depress the sensitivity curves of *several* independent mechanisms (channels): the one centred around the fundamental, and the ones centred around the harmonics.

3.2.1.3 Sub-threshold summation

A *sub-threshold* stimulus is one that produces a response which is too weak to reach detection threshold. In a linear system, two sub-threshold stimuli may add together to produce a combination that may be detected. For example, the result of two identical gratings added together in phase is equivalent to increasing the contrast of the grating from a *sub-threshold* to a *threshold* level.

Sachs, Nachmias and Robson (1971) determined psychometric functions for simple gratings (sinusoids with only one spatial-frequency component) and for complex gratings (consisting of the sum of two sinusoids, each with its own spatial-frequency). Observers were given the task of distinguishing the presence of each grating from a uniform luminance field.

Even though, at the time of Sachs et al.'s writing, there were already a number of evidences that pointed to the existence of *multiple* channels as opposed to a *single* channel, they tested this themselves by formulating two different hypotheses, the "multiple channels hypothesis" and the "single channel hypothesis", which they examined separately. Within the "multiple channels hypothesis", they tested an "independent channels hypothesis" – the scenario that the two components of a complex grating are detected *independently* (i.e. each by a different channel, with the operation of one channel not affecting that of another), and the opposite – the "dependent channels hypothesis".

For the complex grating cases, Sachs et al. (1971) were only able to reject the independence hypothesis when the frequencies of the two components of the complex grating were very close to one another (i.e. a ratio close to 1). For all other ratios, their results indicated *linear and independent channels*, a result which was going to be confirmed by future research (Albrecht & De Valois, 1981; Derrington & Henning, 1989; Henning, Hertz, & Broadbent, 1975). Typically, a ratio of 2 (i.e. a one *octave* difference) between the two frequencies was deemed enough for the components to be detected each by a different channel, with larger separations only being necessary for low spatial-frequencies; this typical octave difference is indicative of the bandwidth of each channel (but see section 3.2.2 for a detailed discussion of this topic).

The single-channel model that they tested in parallel with the multiple-channel model could not at all account for their findings (although they did not rule out the possibility of another, more complex, single-channel model being able to do so).

Sachs et al.'s (1971) results, which show that sub-threshold summation only occurs between complex (compound) gratings whose components are *close* in spatial-

[13] It would have been pointless to look beyond the third harmonic, as at those high spatial-frequencies, sensitivity is much too low to produce any response.

frequency, support the multiple channel model, and furthermore they support the argument that those channels are *independent*.

A more subtle argument has recently been made by Taylor, Bennett and Sekuler (2006), regarding the sub-threshold summation capabilities of the spatial-frequency-selective channels in the presence of visual noise. Taylor et al. measured detection thresholds for several types of broad-band noise and concluded that narrow-band channels can act like broad-band channels in the sense that they can sum information over a broad range of noise frequencies.

3.2.1.4Spatial-frequency-selective masking

Any visual stimulus – natural or artificial – can be seen as consisting of a useful signal (or target) plus noise; the total stimulus is what is obtained if we add together (overlap) the target and the noise.

We have to distinguish between *internal noise* – noise that appears inside the visual system and is a pervasive error which limits detection performance in any task (Burgess, 1985; Patterson & Henning, 1977) – and *external noise*, which resides in the stimulus fed to the visual system.

Internal noise is a low-level attribute of the visual system, and has a number of possible causes. For example, in vision, Burgess (1985, p. 1502) suggests that it is due to "neural noise and random variation of any sort during image data acquisition", while in hearing, Henning (1969) suggests that it is due to variability in the centre frequency of the auditory filter, as well as neural noise. A simple way of quantifying internal noise is through the notion of *equivalent internal noise*, which is equal to the amount of external noise that, if added to the image, would lead to the same degradation in detection performance (Pelli, 1985).

External noise is noise that an experimenter knowingly introduces into a stimulus as a "masker" which is added to the target; the masker is then modulated in order to find the threshold point. It is primarily external noise that is of interest to this section (and, in general, to this thesis).

The degree to which each component of the stimulus (*target* and *external noise*) is present in the total stimulus determines the signal-to-noise (SNR) ratio of the stimulus. If spatial-frequency-selective channels exist in the visual system, a channel will only take in noise that is inside its passband, and reject that which is outside of it. The threshold contrast value can be seen as having a correspondent in a critical SNR value, as measured at the output of the filter. In other words, a certain SNR value is required in order for a *contrast* value to be a *threshold* value.

Stromeyer III and Julesz (1972) studied spatial-frequency selectivity by measuring the threshold for detecting vertical gratings embedded in masking noise consisting of vertical stripes of varying spatial-frequencies. Thresholds were measured both with and without the masking noise present, in order to determine the degree (measured by the *relative threshold elevation*) to which the noise was masking the target grating. As part of

a control experiment, they found that widening the noise band increased the average noise contrast, which, in turn, increased the relative threshold elevation.

In a subsequent experimental condition, gratings of various spatial-frequencies were placed at the geometrical centre of a band of noise (bandpassed noise), and the width of the band was varied symmetrically around that centre. The contrast of the target embedded in noise was set to threshold and masking was measured as a function of the noise bandwidth. It was found that, while masking *did* increase as the noise band was widened, it only did so up to a point: if the noise band was widened beyond a critical width of approximately ±1 octave, masking did not increase further (it actually even decreased by a small amount).

The main conclusion of Stromeyer III and Julesz's experiment is that, if there were only one broad filter spanning the whole range of visible spatial-frequencies, it would be reasonable to expect that masking should increase monotonically for as long as we widen the noise band. Even though at the time, evidence of selective channels already existed, this finding came as an extra proof for the existence, in the visual system, of spatial-frequency-selective channels.

3.2.2 Number of channels and their bandwidth

The paper that first introduced the idea of multiple spatial-frequency selective channels (F. W. Campbell & Robson, 1968) also gives a numerical estimate of the bandwidth of those channels (p. 565): "functionally separate mechanisms in the visual nervous system each responding maximally at some particular spatial-frequency and hardly at all at spatial-frequencies differing by a factor of two". This suggests a bandwidth of about one octave, a value which was also to be found in other subsequent investigations, such as (Burgess, 1985; Kulikowski & Robson, 1999; Stromeyer III & Julesz, 1972).

The human visual system is sensitive to a wide range of spatial-frequencies, from around 0.25 cpd up to 60 cpd (Smallman, MacLeod, He, & Kentridge, 1996). Given the above channel bandwidth estimate, the number of channels can then be inferred to be around eight (also see abstract of (Kulikowski & Robson, 1999))

3.3. Physiological basis for the frequency-selective channels in vision and hearing

It is commonly assumed in the literature that the spatial-frequency channels are hard-wired at an early stage of the visual system, and several studies (Blakemore & Campbell, 1969; Braddick, Campbell, & Atkinson, 1978; De Valois & De Valois, 1980; Robson, 1975; Sekuler, 1974) have pointed to the *striate cortex* (or: *visual cortex*, or: *V1*) as the site of the neurones selectively sensitive to spatial-frequency in the visual system. Evidence for this is the interocular transfer of the spatial adaptation effect discovered by Blakemore and Campbell (1969), as well as the fact that most cortical cells have both *orientational* selectivity (F. W. Campbell & Kulikowski, 1966) and *spatial-frequency* selectivity (F. W. Campbell, et al., 1969).

Campbell and Robson (1968) on the other hand, suggest the locus of the "Fourier transformer" responsible for spatial-frequency selectivity to be in the earlier ganglion cells of the *retina*, although they do not rule out later processing stages in the nervous system also playing a role.

In hearing, according to Plaisted, Saksida, Alcantara and Weisblatt (2003), the site where decomposition of complex sounds into their frequency components takes place is the *basilar membrane*, which contains all of the frequency-selective auditory filters.

3.4. Channels in complex visual tasks

3.4.1 Off-frequency looking

The basic assumption, under the channels model, is that when we look at a *simple grating* (target), we use the filter that is centred around the target's spatial-frequency. However, there are situations in which the channel centred around the target may not be the best one, in terms of the signal-to-noise ratio (SNR) that it offers. This is because the distribution of the noise is such that that particular channel also receives a lot of noise. Therefore, other channels are considered which, even though they take in less *signal* power by being further away from the target, they might also take in less *noise* power, leading to an improved SNR. If after such a calculation, the mechanisms of the visual system decide that the SNR is better than that of the channel which is closer to the target frequency, then the more "remote" channel will be attributed the task of processing the target.

Obviously, when looking at *complex gratings* that have several spatial-frequency components, off-frequency looking might still occur, for each of the individual "targets" represented by the different Fourier components.

An important point to make – one that will become important during the section discussing stimuli – is that off-frequency looking usually occurs if the target is being masked with either *high*- or *low-pass noise*, or even – to a lesser degree – with *bandpass noise*. However, off-frequency looking is unlikely to occur when masking with *notched noise*, because this type of noise prevents the observer from being able to trade off noise *lost* in one channel for (more) noise *gained* in another.

3.4.2 The role of the selective channels in complex identification tasks

The concept of spatial-frequency-selective channels can be used to predict performance with stimuli which are more complex than gratings. For example, Solomon and Pelli (1994) used noise masking (low- and high-pass) to investigate whether complex visual tasks such as reading and object recognition make use of these channels. The premise for this research question was that, while a grating detection task only requires distinguishing the grating from a uniform screen, a task such as letter identification requires a further, more cognitive, process, namely – the classification of the stimulus into one of several learned categories. Solomon and Pelli investigated whether the channels (which are, arguably, a low-level mechanism) could play a role in a high-level task such as letter identification.

Since letters are more spatially compact and have a broader (spatial-frequency-) spectrum than gratings, Solomon and Pelli's (Solomon & Pelli, 1994) initial assumption was that their identification is mediated by not one but by several selective channels (or, at least, by one that is much more broadly tuned). To their surprise, their measurements

of threshold signal-to-noise ratios for identification of letters embedded in noise showed that it is actually the *same* channel that performs both the *low*-level task of detecting *narrow*-band *gratings* and the *high*-level task of identifying *broad*-band *letters*.

This can be easily demonstrated, for example, by looking at an image that contains letters masked by bandpass noise of varying spatial-frequencies; the letters are more effectively masked by a 3-cycles/letter noise than by noise of a lower or higher spatial-frequency. The contrast sensitivity function that this observation suggests is surprisingly similar to the one used for grating detection (described in section 3.2).

In order to directly compare the two tasks (grating and letter identification), Solomon and Pelli used high- and low-pass noise-masking again, this time with the intention of measuring the filter used by the observer to identify the orientation of a tilted grating. The fact that this particular noise masking paradigm (high- and low-pass) was used enabled observers to make use of *off-frequency looking* – a phenomenon which was manifested to the *same* degree in the letter identification task as was in the grating identification task, again leading Solomon and Pelli to conclude that the two tasks must both be mediated by the same filter.

Solomon and Pelli obtained estimates for the filters mediating the two different types of identification tasks (letter and grating), and found them to be very similar. By analogy with off-frequency looking, it could be argued that the observer would choose a *high*-frequency filter to avoid *low*-pass noise and vice-versa, however the two filters (low- and high-pass) were both found to be centred around the spatial-frequency that led to maximum masking in the task described above: 3 cycles per letter. This suggested that filters for different ranges of spatial-frequency are less efficient at the task of letter identification.

Finally, Solomon and Pelli (1994) made a direct comparison between the performance of human observer and that of a particular instance of ideal observer, modelled as a *white-noise ideal classifier*. This model, when given a letter identification task in the context of a noise masking paradigm, would choose the letter that minimizes the total squared difference between the letter and the stimulus. Both the human observers and the ideal observer performed the same task and received the same input, only differently presented: the human observers saw visual stimuli, while the function modelling the ideal observer was given as numerical information corresponding to those stimuli. The results of the performances of the two types of observers were used to derive the shapes of two filters – a *low-pass* filter for the ideal observer and a *band-pass filter* for the human observer.

What Solomon and Pelli managed to emphasise with this study is that a key characteristic of the human perception is significantly different from an idealised observer's: while the white noise ideal derived filter is *low-pass*, the human filter is *bandpass*, which shows that the spatial-frequency-selective channels (which are band-pass filters) are inherent to the visual system, and not just a side-effect of using a very particular and abstract type of stimulus (i.e. gratings) to measure properties of the visual system.

3.4.3 The detectability of spatially- and temporally-varying stimuli; the pedestal effect

The detectability of a periodical signal that is either *spatially*-varying (i.e. sine grating) or *temporally*-varying (i.e. flickering light) can be measured in either *detection* or *discrimination* experiments, usually as part of a 2AFC design. In both categories, the observers' task is to detect in which of the two intervals the target signal was present, with the contents of each interval depending upon the type of experiment:

- in the first category (*detection*), one interval contains a uniform field (i.e. "nothing") and the other contains the target signal;
- in the second category (*discrimination*), both intervals contain a non-uniform visual signal, but the *target* signal is only part of *one* of the intervals.

The design of a discrimination experiment can be so that each interval contains a background in the form of a stimulus having the same frequency (temporal or spatial), orientation and phase as the target, but a different strength (contrast); that background stimulus is referred to as a pedestal[14], because any signal that is used in the intervals always "stands on top" (i.e. is added) to that pedestal.

There are two a priori expectations one can have from a pedestal experiment:

1) One that is admittedly simplistic – that, from the observer's point of view, the pedestals found in both intervals will just cancel each other out, like equal terms found on opposite sides of the equal sign in an equation, transforming the task from a *discrimination* task to an equivalent *detection* task that compares the target alone to an empty (i.e. uniform) field:

$$(\cancel{\text{pedestal}} + \text{target}) \text{ vs } (\cancel{\text{pedestal}})$$

$$\Leftrightarrow (\text{target}) \text{ vs } ()$$

2) One that takes into account Weber's law[15] – that as the background (pedestal) present in each interval increases in strength, a higher strength of the target signal is required for its identification; in other words, that the adding of a pedestal to each interval (i.e. going from a detection to a discrimination task) makes the task more difficult. This would mean that the curve relating target threshold to pedestal strength (known as a **TvC curve**) is a monotonically increasing linear function, as indicated by Weber's law.

Contrary to both these expectations, however, came the results of a series of discrimination experiments done a few decades ago in *spatial* vision (F. W. Campbell & Kulikowski, 1966) and in *temporal* vision (Cornsweet & Pinsker, 1965; Leshowitz, Taub, & Raab, 1968), which emphasised for the first time the existence of an effect – the *pedestal effect* or the *dipper effect* – which results in the TvC curve having a decreasing section (in

[14] Sometimes the two terms "discrimination experiment" and "pedestal experiment" are used synonymously.

[15] Weber's law states that "In order that the intensity of a sensation may increase in arithmetical progression, the stimulus must increase in geometrical progression". This can be expressed mathematically as: "the ratio between ΔS and S – where ΔS is the just-perceivable-difference in stimulus strength and S is the background stimulus strength – is constant".

the shape of a "dipper") before rising in accordance with Weber's law, as discussed above.

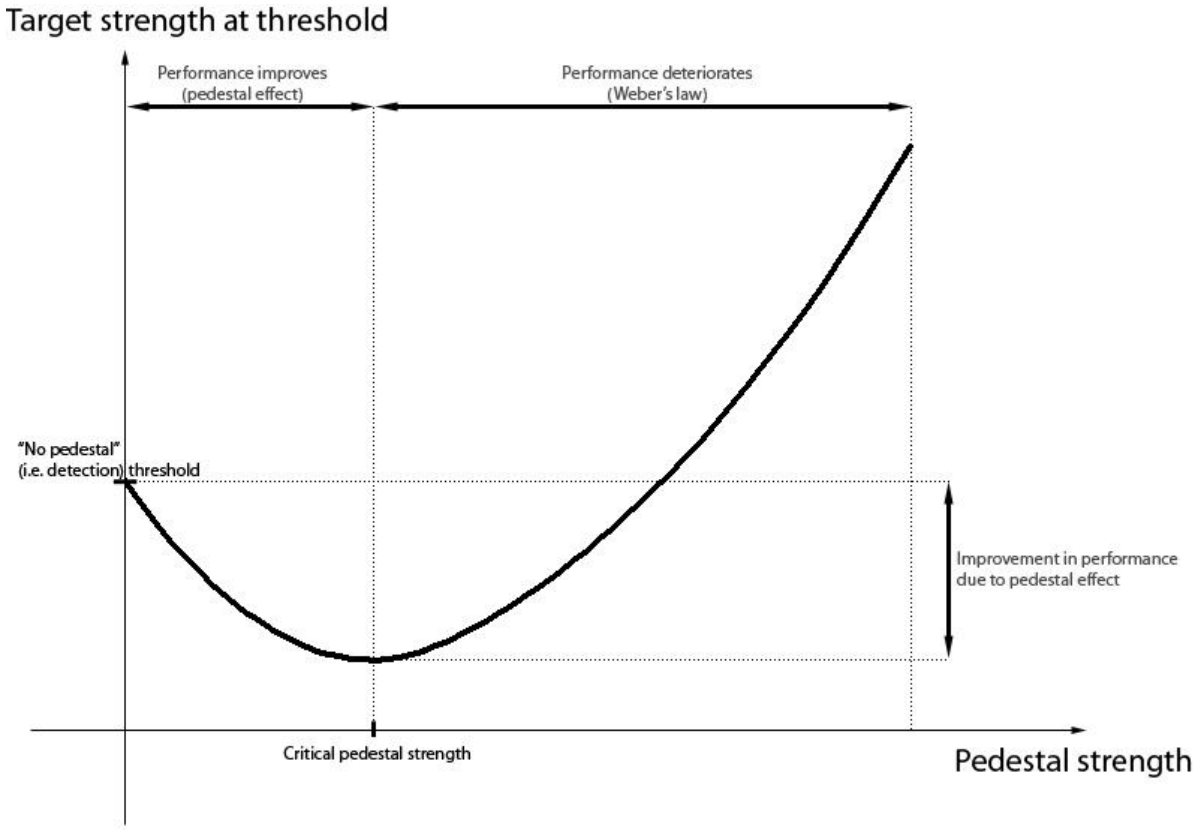

Figure 10: A typical TvC curve that represents target threshold as a function of pedestal strength. In this qualitative representation, there is no noise masker used, only the pedestal masker.
The first (descending) area of the curve represents the performance enhancement that the pedestal brings, and the second (ascending) part of the curve represents the deterioration of performance, in accordance with Weber's law.

For the first part of the dipper, the pedestal is weak enough that it is not seen on its own, but it becomes visible in the other interval, where it is presented together with the target. Even though the presence of the pedestal is considered to be a form of *masking*, the pedestal effectively *helps* the target be seen, which essentially explains the pedestal effect. Thus, the threshold measurements corresponding to the first part of the TvC curve are *detection*-like measurements, because even though the non-signal interval is not blank, the pedestal in it is not seen (it is still too weak). It is when the pedestal grows in strength that it starts to become seen, the task becomes a genuine *discrimination* task, and the pedestal effect starts to disappear (i.e. the curve starts to follow Weber's law).

In order to have a complete model of spatial vision, properties of the spatial-frequency-selective channels need to be known not only in the region of threshold (i.e. properties such as *filter shape* and *filter bandwidth*) but also above the threshold level, i.e. suprathreshold properties. Knowing more about the pedestal effect has allowed researchers to investigate in more depth the suprathreshold properties of the spatial-frequency-selective channels and the mechanisms of contrast gain control which are believed to operate within those channels. This is because including a pedestal in each

interval rises the background contrast to a level which is above that required for the detection of a target signal alone.

Ever since the existence of the pedestal effect was first emphasised, there has been a great deal of debate over the nature and cause of this effect. Several explanations have been suggested:

1. *off-frequency looking*: the pedestal effect is a result of the observer's use of information derived from channels centred around spatial-frequencies other than that of the signal and pedestal;
2. *non-linear transducer*: the pedestal effect is caused by a single mechanism (or: channel), that can be described by a non-linear "transducer"[16] function. This explanation does not necessarily contradict the multiple spatial-frequency-selective channels model, it just suggests that a single (different type of) channel is responsible for the pedestal effect (Foley & Legge, 1981; Legge & Foley, 1980; Nachmias & Sansbury, 1974);
3. *contrast gain control*: the pedestal effect is due to contrast gain control mechanisms, thought to operate at suprathreshold contrasts, which normalise the contrast of input stimuli and integrate them across a broad range of spatial-frequencies ;
4. *uncertainty reduction*: the pedestal effect occurs because the presence of the pedestal in each interval decreases the observer's uncertainty about the target stimulus and makes it easier to identify.

Even in the present, however, there is no consensus about which of those four possible explanations (if any) is the exact cause of the pedestal effect, although a theory starts to emerge that the causes are different for when the stimuli (target and pedestal) are *spatially*- or *temporally*-varying. Two recent papers that investigated the pedestal effect have shed a bit more light onto this: in spatial vision, that of Henning and Wichmann (2007), and in temporal vision, that of Smithson, Henning, MacLeod and Stockman (2009, under review).

Henning and Wichmann (2007) used a sinusoidal grating as a pedestal, which had the same spatial-frequency, orientation, phase and duration of presentation (in seconds) as the target grating to be detected. A single spatial-frequency was used constantly, both for the target and the pedestal. Two types of maskers were used, as described above (pedestal and noise), with the noise's notch width and the pedestal's contrast being manipulated experimentally. Just as the noise could be absent from certain conditions, the pedestal's contrast could also be reduced to zero, such that the experiment became a *detection* rather than a *discrimination* experiment (see explanation above).

Henning and Wichmann found that the pedestal effect was

- *very large*, when not using a noise masker
- *large*, when masking with *broad-band noise*

[16] A transducer is a device that converts one type of energy into another – for example, a loudspeaker or an electrical sensor. In this context however, by "non-linear transducer" we understand a mechanism that produces an output that is a non-linear function of the input.

- *very small*, when masking with *notched noise*

and since notched noise prevents the employment of channels not tuned to the pedestal frequency (i.e. prevents off-frequency looking), they concluded that *off-frequency looking* must be the real explanation of the pedestal effect. Although Henning and Wichmann do not claim their result also extends to temporal vision, neither do they admit that there, the other models might better explain experimental data than the off-frequency looking model.

Smithson et al. (2009, under review) also used masking with pedestals and (notched) noise in detection and discrimination experiments that investigate the mechanisms underlying flicker perception, where the pedestal effect again plays an important role. While Henning and Wichmann found that, in *spatial* vision, the dipper effect is almost unnoticeable when using notched noise, Smithson et al. noticed that the pedestal effect emphasised using *temporally*-varying stimuli *does* survive even in notched masking noise. Since notched-noise masking makes it difficult for off-frequency looking to occur, they interpreted this as evidence that models based on off-frequency looking (i.e. models that assume observers employ channels tuned to several temporal frequencies), cannot explain the pedestal effect, at least in temporal vision. In other words, the pedestal effect could very well exist even if a *single* channel was employed.

The only model that could account for the entirety of Smithson et al.'s data set was the *non-linear transducer model*, which claims that the pedestal effect is due to a single mechanism, characterised by a specific non-linear transducer function, the expression for which Smithson et al. have derived.

From the vantage point of my thesis, the idea that the pedestal effect (i.e. the enhancement in performance) in spatial vision might come about due to a flexible combination of multiple channels (i.e. off-frequency looking, as suggested by Henning and Wichmann) is interesting, and something to which I will return.

The pedestal effect has also been shown to exist in hearing (Pfafflin & Mathews, 1961; Raab, Osman, & Rich, 1963), which is not surprising given that *amplitude* discrimination of an auditory stimulus is formally equivalent to *contrast* discrimination of a visual stimulus.

3.5. Auditory frequency

In hearing, *frequency* (*temporal-frequency* or *auditory-frequency*) refers to the number of *cycles per second* of the air pressure waveform. In vision, *spatial-frequency* measures the number of *cycles per degree of visual angle* that a certain image exhibits.

I have looked at how, in vision, different spatial-frequencies are processed by individual spatial-frequency-selective channels. Similarly, in hearing, the different auditory-frequencies are processed by individual, frequency-selective channels.

3.5.1 Off-frequency listening

In section 3.4.1, I have defined the phenomenon of *off-frequency looking*, which may occur when an observer is presented with a visual stimulus that is masked with

certain types of noise. Because of the similarity of the channels model in the visual and auditory case, a similar phenomenon exists in hearing, where off-frequency *listening* refers to the auditory system's "choice" to employ a channel other than the one closest to the signal's frequency; this choice is, again, based on a direct comparison between the various channels' SNRs.

Since many experiments in both vision and hearing intend to study properties of specific channels, it is often not desirable that off-frequency looking/listening occur, and to ensure it does not, experimenters often use *notched* noise in their stimuli, to mask their targets. One example – in hearing – is the method used by Patterson (1976), which ensures that the observer always "listens" through the auditory filter centred around the signal frequency. It is for the very same reason that I too have used *notched* noise in my experiments, as opposed to other types of noise filtering.

3.6. Auditory filters

In the history of this type of investigation, two different approaches have produced different (yet sometimes surprisingly converging) results: neurophysiological and psychophysical measurements. While the former approach addresses single units such as individual cochlear fibres, the latter measures the overall properties of the nervous system as a whole, effectively treating everything between the input (i.e. stimulus) and output (i.e. response) as a "black box".

The earliest estimates of auditory filter bandwidths come from psychophysics experiments done on *human* observers in the 1950s (Zwicker, Flottorp, & Stevens, 1957), which estimated the filter responsible for detecting 1 kHz tones (or rather: the one assumed, as a result of assumptions regarding off-frequency listening, to operate in response to detecting a 1kHz tone) to be about 160 Hz wide. A similar figure was reported a decade later (de Boer, 1967) by a measurement of the same filter's bandwidth, only this time done upon a *single cochlear fibre* in the *cat*. Even though the cochlea in the cat is not the same size as the cochlea in the human, the fact that a neurophysiological measurement could even partly account for psychophysical data was a first indication that the frequency selectivity available at the output of the cochlea (in either cat or human) may not necessarily be increased if we look beyond the cochlear fibre itself. This theory was apparently proved wrong, then confirmed and then again proved wrong by several psychophysics experiments done on humans in the mid '70s which used noise masking. I will discuss those papers in the next few paragraphs.

The first difficulty in relating psychophysical performance to underlying channels is the possibility that observers would use a filter that is not tuned around the target signal, regardless of the type of noise that the target is embedded in. This is the issue I identified in section 3.4.1, termed off-frequency looking. The auditory counterpart is off-frequency listening (section 3.5.1).

The second difficulty is that, when trying to estimate a property of a frequency-selective neural filter, such as its attenuation characteristic, a mathematical model is required to relate threshold data to a numerical evaluation of the filter's shape (e.g. the width of its passband).

The first paper – in a historical sequence – was (Houtgast, 1974), which measured the detection threshold of a tone masked with band-passed rippled noise (=noise whose spectrum varies sinusoidally). Houtgast assumed that, despite the presence of the masker, the auditory filter would still be centred around the tone frequency, because this is how the best signal-to-noise ratio would presumably be achieved. In other words, he assumed no off-frequency listening takes place that would shift the filter off the target frequency, since doing so would offer no improvement in SNR. This assumption, coupled with the one that the entire filter shape can be approximated by a Gaussian curve, led Houtgast to a bandwidth value of 170 Hz (for the filter centred around a target tone of 1 kHz). Again, a *psychophysics* experiment done on *humans* had produced the same bandwidth estimate as a *neurophysiological* measurement on *cats* (de Boer's study).

The three papers mentioned so far in this section ((Zwicker, et al., 1957), (de Boer, 1967) and (Houtgast, 1974)) all seem to suggest that the frequency selectivity of the auditory system (i.e. the sharpness of the auditory filters) is determined early, in the cochlea, rather than in loci situated in further stages of neural processing. Three other papers, which I will discuss in the paragraphs to come, have, however, found – using *psychophysical* measurements – narrower bandwidths for those filters, suggesting that a much sharper auditory filter is in fact available than would be indicated just by *neurophysiological* data.

In an experiment similar to Houtgast's, Patterson (1974) measured detection threshold for a tone embedded in low-pass noise, in terms of the cut-off frequency of the noise. Patterson obtained, for the 1.0 kHz filter, a bandwidth of 59 Hz – much narrower than Houtgast did.

A second low-pass noise masking experiment, which confirmed Patterson's (1974) low bandwidth estimate was that of Margolis and Small (1975), which differed from Patterson's in the way that the noise was generated (digitally); this led to a much sharper noise edge and provided a better approximation to a step function than the one Patterson (1974) used. Margolis and Small (1975) obtained the filter shapes in the same way that Patterson did – by differentiating the curve that related threshold to the cut-off frequency of the low-passed noise. Their filters were quite similar to those reported by Patterson (e.g. 57 Hz passband for f_0 = 1 kHz), only slightly less symmetric.

Patterson's (1974), as well as Margolis and Small's (1975) narrow bandwidths estimates, seem to prove a point which is the exact opposite of that made by the previous set of three papers: that neural stages of the auditory system beyond the cochlea seem to actually *do* play a role in the sharpening of the filter centred around 1 kHz at least under some circumstances. It should be noted, however, that both (Patterson, 1974) and (Margolis & Small, 1975) made the same assumption about off-frequency listening *not* occurring, an assumption that may be invalid, and that could have led to underestimation of filter bandwidth.

Patterson later (1976) redid the experiment, assuming off-frequency listening *does* occur – but he used a *notched* (rather than a low-passed) noise masker, in order to make this strategy non-optimal and thus unlikely to occur.

Once again, by using a notched noise masker, if the filter is reasonably symmetric then it can be assumed to be centred at a point near the tone, because this will be the region where the SNR at the output of the filter is greatest. If the filter were shifted up in frequency, then the total noise power at the output of the filter would be increased and

the power of the tone would be decreased, which would, overall, decrease SNR – therefore this alternative can be ruled out when using notched noise.

In addition to this assumption, Patterson (1976) only used the Gaussian approximation for the filter's shape in the region of the *passband*, rather than for the *entire shape* of the filter.

Since the discussion of the six classic investigations into the bandwidths of the auditory channels done so far was rather complex and may have confused the reader, I have synthesised all of their results in Table 2 below:

Table 2: Comparative results of classical neurophysiological and psychophysical investigations into the bandwidths of auditory filters.

When	Who	How	On whom	passband for f_0 = 1 kHz	passband for f_0 = 3 kHz
1957	Zwicker	psychophysics: Gaussian approximation for the *entire* filter shape	humans	160 Hz	500 Hz
1967	de Boer	neurophysiological measurements	cats	160 Hz	750 Hz
1974	Houtgast	psychophysics: band passed rippled noise, Gaussian approximation for the *entire* filter shape, filter assumed to be centred around the tone	humans	170 Hz	-
1974	Patterson	psychophysics: low-passed noise, filter assumed to be centred around the tone	humans	59 Hz	-
1975	Margolis and Small	psychophysics: low-passed noise, noise generated digitally, filter assumed to be centred around the tone	humans	57 Hz	-
1976	Patterson	psychophysics: notched noise, Gaussian approximation for filter's shape in the region of the band pass, filter *made* to be centred around the tone by the use of notched-noise masking	humans	140 Hz	-

The bandwidths obtained by Patterson (1976) for the three filters (centred around 0.5 kHz, 1 kHz and 2 kHz) are more than double those reported by (Patterson, 1974) for the same three tone frequencies, which suggests that the auditory filter was not centred around the tone in (Patterson, 1974), i.e. off-frequency listening *did* occur.

The result of Patterson's (1976) paper leads once again to the conclusion that the frequency selectivity manifested at the output of the cochlea is consistent with auditory filter shapes derived with noise-masking experiments. Those filters are well-tuned, although they are more than twice as wide as those reported in the previous two papers ((Patterson, 1974), (Margolis & Small, 1975)). This implies that in those two experiments, the auditory filter was not centred around the tone (as was assumed), and therefore what was measured was actually the shape of the *skirts* of the filter rather than the shape of its *passband*, as was intended.

Later research has confirmed Patterson's final result, by producing bandwidth estimates[17] for the auditory filters that were roughly 10-12% of the filter's centre frequency (Moore & Glasberg, 1981), with filters becoming slightly wider at higher frequencies and/or for higher levels of the stimuli used to derive them (Glasberg & Moore, 1990).

[17] Filter bandwidths are usually given as ERB (equivalent rectangular bandwidth), which are obtained by modelling the filters as ideal (rectangular) band-pass filters.

4. PRIOR KNOWLEDGE AND PERCEPTUAL LEARNING

The channels model is well accepted both in the visual and auditory psychophysical literature, and I will consider it to be true for the purpose of this work. However, I am also challenging the channels model in a way, and aim to extend it, as will be further described.

Up until now, there have been few studies that looked into what happens if those channels are used in a repeated context. If the channels were truly hard-wired – indeed *filters* in the engineering sense of the word – then one would expect no change due to practice in a repeated context.

I suspect that a change might, however, take place that has to do with the filters – and it might either be a characteristic of the filters themselves (e.g. they might get narrower), or it might be the way that a later system combines information from the filters (e.g. off-frequency looking and how we can attend to several filters at once).

In the following sections, I will summarise previous work that has investigated specific forms under which perceptual learning can occur due to repeated exposure to visual stimuli with certain predefined characteristics.

4.1. Use of prior signal knowledge (the SKE condition)

Uncertainty regarding various signal parameters (related, for example, to temporal presentation, spatial location or phase) has the potential to decrease an observer's performance in a visual or auditory detection task. This fact is both predicted by signal detection theory (see section 2.1) and pointed out experimentally (Pelli, 1981). It therefore seems reasonable to assume that, in a visual detection task where stimuli are embedded in noise, if an observer has prior knowledge about the upcoming signal to be detected, then he will not be misled by noise which looks different from the "known" signal, and thus he will be able to improve his performance.

Nevertheless, this was a contentious topic at the time that Burgess (1985) wrote his seminal paper which investigated this very question, and demonstrated

unambiguously that this is indeed the case by obtaining experimental results which were very different in the presence vs in the absence of prior knowledge.

Burgess proposed that the way the human visual system makes use of prior knowledge is by a strategy involving cross-correlation between the signals that are known to be possible targets and the given signal, followed by a selection of the signal that has the highest a posteriori probability according to Bayes' theorem[18]. This is the same strategy that an ideal observer would use, only done "suboptimally".

Signal detection theory can only be used to describe the *ideal* observer, and not a *human* observer. Nevertheless, it is useful because it gives researchers a theoretical limit with which they can compare the performances of human observers that were studied experimentally. The particular comparison that was of interest for Burgess was between human and ideal performances under two distinct experimental conditions: *signal known exactly* (SKE), where the observer knew precisely which signal he was expected to detect, and *signal uncertainty*, where the observer did not know which of the possible set of signals he was expected to detect.

The two conditions were used in two types of tasks:

1. 2AFC *detection* tasks, which required observers to select which of the *two* noise fields presented had a signal embedded in it; the signal could be known (*SKE* condition) or unknown (*uncertainty* condition);
2. 10AFC *discrimination* (or: *identification*) tasks, which were always done in the *uncertainty* condition and which required observers to identify which of the possible signals was embedded in the (*only*) noise field presented.

Initially, a set of 16 orthogonal signals were considered as targets, but for the purpose of having homogenous detectabilities for all visual targets, a preliminary experiment (2AFC task, under the SKE condition) was run which enabled the selection of only the 10 most detectable[19] signals of the set to be used in the main experiment. The orthogonality of the signals was important, because it ensured that all possible decisions that the observer can take are statistically independent – a condition which is necessary in order to be able to analyse an observer's performance.

Burgess chose to mask the visual targets with noise because this allowed the experiment to be done with *suprathreshold* stimuli, which correspond more closely to normal visual conditions than close-to-threshold stimuli. Furthermore, noise masking constituted a performance limitation, and thus allowed direct task performance comparisons between human and ideal observers.

Plots of the detectability index, d', and of the equivalent measure of percentage of correct responses were generated, as a function of the signal SNR, in all possible combinations: for the *ideal* and *human* observers, in *2AFC* and *10AFC* tasks, and under the *SKE* and *uncertainty* conditions.

The results for human observers under the SKE condition indicated performance superior to that of the ideal observer, for low SNR values. The results were also relatively linear, and since visual detection tasks that use noiseless stimuli generally produce non-

[18] In probability theory, Bayes' theorem is often used to compute a posteriori probabilities given a priori observations. The theorem is able to explain how the probability that a theory is true is affected by a new piece of evidence (Knill, Friedman, & Geisler, 2003).

[19] In terms of the value of the d' parameter (see section 2.1. on SDT).

linear results, this can be taken to mean that, whatever we deem to be the source of human inefficiency in the SKE condition, the imperfect human performance is not due to signal uncertainty, but rather can be attributed to factors such as *internal noise*.

The detectability index d' for the *uncertainty* condition was plotted as a function of the same parameter d' but for the *SKE* condition. The curve describing the ideal observer provided a reasonable fit to the data collected for the human observer, which lead Burgess (1985) to conclude that the same source of human inefficiency must be present for both the *SKE* and the *uncertainty* condition.

The overall conclusion, drawn from the pronounced difference in results for the *SKE* vs the *uncertainty* conditions, was that humans clearly can use prior information in visual signal-detection tasks.

At the time of the writing of this Burgess (1985) paper, two diverging schools of thought existed in relation to the role that prior knowledge plays in human visual perception:

1) the first is the *probability summation model*, which did not provide a good fit for Burgess' data. Probability summation simply means that an observer will say "yes" to a detection task if any of several events occur. According to the probability summation model, prior knowledge is not used during a visual signal detection task. Instead, detection probabilities are summed together over a set of parallel channels, an operation equivalent to a series of cross-correlations between the visual input and each visual filter. According to the model, detection occurs if the output of each cross-correlation exceeds a fixed threshold.
2) the opposite view – that prior knowledge *does* play a role – is consistent not only with Burgess' results but also with previous findings in Artificial Intelligence ((Marr, 1983), cited in (Burgess, 1985)), which suggest that vision requires *low*-level operations (such as data acquisition and internal representation) as well as *high*-level operations (i.e. interpretation of the low-level data according to an observer's knowledge about the physical world).

Two papers co-authored by Burgess and Ghahdeharian (1984a, 1984b), which preceded (Burgess, 1985), had attempted to assess human detection and discrimination performance with noisy stimuli, and had found this performance to be close to that of an ideal observer similarly described to the one in (Burgess, 1985). This ideal observer is assumed to only be limited by stimulus *noise* and *uncertainty*, and it is for this reason that these two factors played in important role in (Burgess, 1985), which followed up on (Burgess & Ghandeharian, 1984a) and (Burgess & Ghandeharian, 1984b).

In the decade following Burgess' papers, a series of studies on the mechanisms of perceptual learning, done by Dosher and Lu, further pointed out the ability of the human visual system to make use of prior knowledge and experience. I will describe those studies in the following section.

4.2. Perceptual learning and flexible channels

4.2.1 Channel reweighting

Observers' performance in perceptual tasks (such as a visual discrimination task) often improves with training – this is what defines *perceptual learning*[20]. However, to date, the mechanisms that underlie perceptual learning are not fully understood, although one particular model stands out which is increasingly agreed upon. This is the *channel reweighting model*, towards which I will lead the development of this section.

Accepting this model and using an external noise masking paradigm offers a complete method of testing different hypotheses regarding the specific submechanisms of perceptual learning. This is the case in the two Dosher and Lu papers (1998, 1999) that I will discuss further on in this section.

The notion of *transfer of learning* refers to the ability of an observer to transfer the learning developed in one context to another context that shares a number of characteristics with the first one. The context can relate to the given task, to a certain target or to a certain mechanism involved in performing the task. For example, *interocular transfer* refers to the ability of an increase in performance due to learning a visual task performed with one eye, to be transfered to the other eye.

Mollon and Danilova (1996) have pointed out that perceptual learning is *stimulus-specific*, in the sense that if a visual stimulus is slightly changed in appearance (e.g. it is tilted) or in the way it is presented (e.g. to one eye or the other), then after repeated trials, observers show much less transfer of learning to the changed stimulus than they do if the stimulus had not been changed. This could be interpreted in more ways than one: for example, it could be taken to mean that learning is a low-level mechanism, but – as Mollon and Danilova point out – it might equally just mean that the plasticity of the learning implies a change taking place not in the (low-level) channels themselves but in the way more central mechanisms are able to discern which channel outputs are more useful for the given task.

In other words, the "specificity" of perceptual learning might have to do with what is learnt, rather than with who (what mechanism) is doing the learning. The literature did contain, at the time of Mollon and Danilova's writing, arguments in favour of learning that is specific to the retinal location: for instance, in Karni and Sagi's (1991, 1993) visual discrimination perceptual tasks, the only improvement that was observed was specific to the retinal position that had been trained, and there was no transfer of learning to untrained locations in the retina. Newer papers have, however, started challenging the idea that perceptual learning necessarily implies specificity – for example, it has been shown that in the context of a Vernier discrimination task, learning transfers completely across retinal locations (Dwyer, 2008).

The exact nature of the plasticity of perceptual learning was investigated for the first time by Dosher and Lu (1998), who contrasted two possible mechanisms of learning:

[20] One should not confuse perceptual learning, which is low-level, with cognitive (high-level) learning, which is what comes into play when, for instance, someone learns the rules of arithmetic.

selecting (giving different weights to) some channels and not to others, or effectively changing the bandwidth of each channel ("channel fine tuning").

Most studies assume the presence, in the visual system, of *internal noise*, which had been shown to deteriorate performance in the *absence* of *external* noise (Legge, Kersten, & Burgess, 1987). Dosher and Lu's (1998, 1999) ran experiments involving orientation discrimination tasks in which *external* noise was manipulated experimentally, in order to infer how that external noise might affect perceptual learning. Specifically, they made use of external noise as a masking factor, which they contrasted with the "masking" done by the noise that is *internal* to the observer's own visual system.

Dosher and Lu (1998) defined two different threshold criteria (performance levels), and for each one of them, the curves that represented contrast threshold as a function of the contrast of the external noise manifested two outstanding regularities:

1. for any criterion, the curves representing data collected *late* in the practice were lower (i.e. better performance) than the ones representing data collected *early* in the practice. The greatest improvements were noticed across the first few sittings, suggesting that there is a higher degree of perceptual learning occurring at the *beginning* of an experiment.
2. the ratio between any threshold obtained in a given sitting for the first criterion, and its "counterpart" threshold for the second criterion was always constant. In other words, for any given sitting, the curves representing the two criteria would be parallel to each other, if plotted on the same log contrast scale.

Each curve from the two graphs manifested a "low limb" – for low external noise contrast, and where threshold was relatively constant – and a "high limb" – for high external noise contrast, where threshold was beginning to rise. Dosher and Lu (1998) formulated the following hypotheses in regards to performance within the two noise limbs:

Table 3: Hypotheses formulated by Dosher and Lu (1998) in regards to the causes of improvement and deterioration in performance within the low and the high noise limbs.

The low noise limb	**The high noise limb**
Performance is limited by inefficiencies in the visual system, expressed as *equivalent internal noise*	Performance is limited by *external noise*.
Improvements are due to *stimulus enhancement* (through *reduction of additive internal noise*)	Improvements are due to *external noise* exclusion.

They tested these hypotheses using a model that describes the observer as an input–output system and the perceptual system as consisting of a combination of

- a perceptual template (or: filter) tuned to the stimulus, which decides a certain configuration of weighted inputs from the visual channels;
- a non-linear transducer function, which reflects non-linearities early in the visual system;

- multiplicative internal noise and additive internal noise, which together characterize processing inefficiencies in the visual system; and
- a statistical decision rule.

This model yielded curves that correctly predicted the observers' performance if the assumption was made that perceptual learning is only due to *one* of the three suspected mechanisms:

1. external noise exclusion, which improves performance for high external noise levels
2. stimulus *enhancement* via reduction of additive internal noise, which improves performance for low external noise levels
3. reduction of multiplicative internal noise.

By matching the curves predicted by the mode to their data, Dosher and Lu (1998) were able to rule out reduction of multiplicative internal noise[21] and thus concluded that perceptual learning must be due to a combination of stimulus *enhancement* and external noise *exclusion*. In other words, there might be two components to perceptual learning: *enhancing the perceptual template* (i.e. learning the target) and making noise properties predictable (i.e. *learning the noise*).

Hurlbert (2000) exemplifies this process by examining how one becomes able to appreciate music played on an old LP that is marred by crackling sounds. The "practice" of repeated hearings triggers perceptual learning: one becomes able to both ignore the noise and more easily follow the musical line that one has come to know over time. This is equivalent to being able to efficiently integrate over relevant information from the signal.

Previous studies had been able, by using discrimination tasks with stimuli located in different parts of the screen (and thus, leading to different retinal positions), to identify the first of the two above-mentioned mechanism – that of *external noise exclusion*. By using external noise as a variable, Dosher and Lu (1998) were able identify the workings of the second of the above-mentioned mechanisms – that of *stimulus enhancement*.

The two simultaneous improvements at the observer level enabled them to suggest that the main function of perceptual learning takes place not by directly *retuning* the basic channels but by (indirectly) placing different levels of importance on each of them, or in other words *reweighting* the channels. Reweighting implies selecting the channel that is most appropriate for the given task, and reducing inputs from all others. The selection of the best channel(s) would then strengthen the connection between its (their) output(s) and a learned categorisation structure.

The channel reweighting model is able to explain – unlike the channel retuning model – why simultaneous learning of motion discrimination tasks can take place for different directions, as shown in (Liu & Vaina, 1998). Furthermore, it could be confirmed by a number of neural network models (such as (Grossberg, 1974) and (Rumelhart & McClelland, 1986), as cited in (Dosher & Lu, 1998)).

One of the most important conclusions that has been drawn about perceptual learning – that it takes place by channel reweighting, which implies central mechanisms

[21] This mechanism implied changes in non-linearity, while the data showed a constant non-linearity factor (γ) throughout.

being able to selectively use outputs from earlier levels – is found to be consistent with newer research that models low-level perceptual learning, such as the study by Petrov, Dosher and Lu (2005). Petrov et al. obtained a model that fitted their orientation discrimination data remarkably well, and which once again validates the multichannel reweighting model, all the while invalidating the "channel retuning" model.

4.2.2 Perceptual learning and context

An interesting link can be made between perceptual learning and the idea of "context" (i.e. having in sight not only the trials themselves but also the pattern in which they are presented). Normally, when perceptual learning effects are investigated, training is done with stimuli that maintain constant parameters throughout the experiment. The result is usually that performance improves with repetition. This holds even for stimuli of multiple parameters, if the particular learning effect being investigated is transferable across stimuli, and if each stimulus parameter is practiced in a separate block (as is usually the case).

However, the idea of context becomes important if one is to present stimuli of different parameters, and present them not in a blocked but in a randomised fashion. Specifically, I am referring to the contrast discrimination experiments performed by Adini, Wilkonsky, Haspel, Tsodyks and Sagi (2004), which have used an experimental manipulation called "contrast roving", whereby the several visual stimuli to be learned – all of different contrasts – are randomly interleaved from trial to trial. This way of manipulating context has led Adini et al. to find poorer performance (i.e. smaller learning effects), which is presumably due to stimulus uncertainty.

So in effect, Adini et al. have shown that "contrast roving" can prevent perceptual learning or that, in other words, perceptual learning does not take place (or at least not to the expected degree) if the presentation of stimuli of different contrasts is temporally shuffled (i.e. trials are randomised). This idea comes in support of the motivation for the current work, which is expanded on in chapter 5.

4.2.3 Entropy masking: the effect of predictability of masks

Traditionally, there have been two types of (visual) masking used in vision experiments: *contrast masking* (used in pedestal-type experiments, as presented in section 3.4.3), which is deterministic and works by decreasing gain in the (early) visual system, and *noise masking*, which is random and works by increasing the variability of internal decision variables. It is not always straightforward to ascribe more complex background stimuli that may be used as maskers of either category. Watson, Borthwick and Taylor (1997) pointed out the need for a model that can explain the variability observed in the threshold-elevating effect of complex maskers such as *bandpass noise* or *natural images*.

Watson et al. (1997) did a series of experiments that measured contrast thresholds for a Gabor target added to a background masker. There were five types of maskers (no mask (a uniform field), a cosine grating, a band-passed noise, a white noise and a natural image), and several possible conditions relating to the presentation of the noise maskers. Three conditions of particular interest to this discussion were: "same", which meant the same sample of noise was added to the two intervals of a trial, "unique",

which meant different samples were added to the two intervals, and "fixed", where a single sample of noise would be used in each interval of every trial.

Watson et al.'s results clearly showed that the type of mask has a big influence on how much it elevates threshold, with bandpass noise achieving the highest threshold elevation, followed by white noise, the natural image, the sine grating and finally the uniform field.

Presentation condition also strongly influenced the amount of threshold elevation, particularly in the case of the "fixed" condition, which showed significant learning: across different estimates, performance decreases from an initial high level comparable to that found for the "unique" condition, to one close to those obtained with a simple sine-wave masker. In other words, having the same sample of noise in each interval of every trial made the task as predictable as if there were only a single, perfectly predictable frequency component (as is the case of the sine grating).

Since both the *type* of the masker and the *condition* it was presented in have been found to contribute to the masker's effectiveness, it could be said that it is the *predictability* of the masker as a whole that is important. This goes against the traditional assumption of the notched-noise paradigm, which is that noise properties cannot be learnt. Watson et al. (1997) argue that our ability to learn noise is related to its "simplicity"; they introduce the term "entropy masking" to indicate that the power of a masking stimulus is related to its unpredictability for the observer

One interesting part of Watson et al.'s results showed that, while the "unique" condition did elevate threshold by a significant amount, it did not do so more than the "same" condition. This was surprising because, even though both conditions have an effect on contrast gain, it is only the "unique" condition that has the random character expected to make learning difficult. Perhaps even more importantly, this result goes against the perceptual template model associated with an ideal observer, according to which, adding the "same" sample of noise to both intervals should not elevate threshold in comparison to a situation in which the noise were absent altogether (because the ideal observer can simply "subtract" that same sample of noise). As a side note, I should mention that so far, only an incompletely specified perceptual template model (i.e. one specified with a single set of parameters) has been found to be able to explain human observers' performance over a wider range of performance levels (Lu & Dosher, 1999, 2001).

It is interesting to note the results obtained in a similar study, done by Ahumada and Beard (1997), which compared all the conditions mentioned previously: the one where the noise samples were the same for each interval of every trial; the one where the two intervals had equal noise samples but the samples were different among trials; and the one where every interval of every trial had a different sample of noise. They found that detection performance was only better in the first condition, but was the same in the second and third conditions.

Watson et al.'s (1997) and Ahumada and Beard's (1997) findings that using the same noise sample on each trial does not lower thresholds compared to having unique samples on each trial had also been obtained by other researchers (Eckstein, Ahumada, & Watson, 1997a; Swift & Smith, 1983). In addition, Eckstein et al. (1997b) also investigated – using a spatial four-alternative forced-choice (4AFC) detection task – the effect of using the same exemplar of noise masker vs using different exemplars. It seemed that in this

case, observers *were* able to use the repeated nature of the masker to improve their performance.

A conclusion that would naturally follow those presented in the Watson et al. (1997) and Eckstein et al. (1997b) papers would be that we adjust our performance based on the noise characteristics.

5. MOTIVATION FOR THE PRESENT STUDY

As discussed in the literature review done in the previous two chapters, it is almost universally assumed, in the research that uses notched-noise masking to estimate channel properties in vision and hearing, that the blocking of different notch-widths is not critically important. Often, such as the case of Baker and Rosen's paper (2006) it is even unclear from the methodological account whether or not different notch-widths were blocked, but it is certainly common to do so. At the same time, I have also reviewed literature that suggests that the predictability of a masking stimulus is important (Eckstein, et al., 1997b; Watson, et al., 1997).

I propose a series of experiments that will emphasise certain characteristics of the *frequency*-selective channels that exist in the human auditory and visual systems. Specifically, I want to determine the minimum contrast needed to accurately discriminate the target stimulus when it is embedded in notched noise, and how this threshold depends on the width of the notch. This is the classic experiment using notched-noise masking to estimate channel properties. However, I additionally want to test whether the results from such an experiment depend upon the way in which trials are presented.

Are estimates of spatial-frequency channel bandwidth derived from classical notched-noise masking experiments dependent on context? In other words, can the performance of an observer in a single trial be considered in isolation, or do we also need to look at the context in which that trial was presented?

I choose to manipulate "context" by measuring thresholds for particular notch widths, either in blocked or randomised presentations. In a blocked session, observers will be exposed to only one notch width for the duration of the session, whereas in a randomised session, all values of notch width will be present as part of the same session.

According to the classical model, it should not matter where exactly the trial is, temporally, among the other trials. If I find that it *does* matter, then perhaps the visual system can learn from the "notch width experience" and, if so, notched-noise masking experiments should not be considered low-level (i.e. a hard-wired property of the visual system), but a more flexible process, like the channel reweighting first identified in (Dosher & Lu, 1998, 1999).

The central point in this thesis is whether context or repeated exposure to a particular noise condition can influence estimates of channel properties (traditionally thought of as hard-wired). In other words, whether "filter" properties might be

dynamically adjusted on the basis of context, rather than being fixed, and completely low-level.

As highlighted in (Dosher & Lu, 1998, 1999), performance might improve for different reasons. One possibility, for example, is that observers might be able to obtain lower thresholds for large notch widths in blocked conditions relative to randomised conditions. This is because if the presentation is blocked and the notch width is large, energy at the target frequency is (a priori) unlikely to come from noise. They could, therefore – in SDT terms – lower their *criterion*[22] (see section 2.1 on Signal Detection Theory) for attributing perceptual activity to "target+noise" rather than to "noise alone". In randomised conditions, even when – by chance – the notch width is large, the same lowering of the criterion is not possible.

There might also be, however, other ways in which observers could adopt a flexible strategy. For example, they might look at the output of several channels, rather than just the one at the target frequency. In previous work, people have talked about this as off-frequency looking (or listening), which typically refers to "looking at" ("listening to") (i.e. using) channels that are tuned *above* or *below* the target frequency. Here, I would be suggesting that observers used channels that were symmetrically placed around the target frequency (since I use notched-noise), but that in blocked conditions with large notch widths, they might use *more* of them, spanning a wider range of frequencies.

The simplest prediction that I can make is, therefore, that the relationship between notch-width and threshold might be *different* for the blocked and randomised conditions – without making any predictions about *how* it might be different. If I were, however, going to give a direction, I would predict that thresholds should be *lower* in the blocked conditions than in the randomised conditions.

To answer these questions, I need to collect data in the form of threshold as a function of notch width, in both blocked and randomised conditions. For consistency with other research that looked at channel bandwidth, I chose to use a two-interval forced-choice task, which was also used by other psychophysical research in vision (for example, by Henning and Wichmann (2007)) and in hearing (for example, by Plaisted et al. (2003)).

One additional manipulation – independent from the first one, but that might emphasise phenomena from the same "realm" – has to do with how the two samples of noise that were added to each interval of the 2AFC trial were related to each other: the "same" condition, in which the same sample of noise was added to the two intervals of a trial, contrasted with "unique", which implied that different samples were added to the two intervals. As summarised above, the ideal observer model suggests that the "same" condition would make noise perfectly predictable, effectively allowing the observer to "subtract" the noise from both intervals and perform the task as if the noise were not present in either of the intervals.

In practice, it has been shown (see section 4.2.3) that thresholds measured in 2AFC trials do not actually always follow this prediction, and performance for the "same" condition is sometimes even worse than in the "unique" condition. I include this manipulation in my study, and additionally look to see whether it interacts with the other

[22] I have used a 2AFC task, so the word "criterion" needs to be taken in a generalised way – one of the arguments for using 2AFC in general is that it encourages observers to set a *neutral* criterion (i.e. to reduce *bias*).

independent variables I use – namely, notch width and "blocked" vs "randomised". One more reason why it is of interest to add the "blocked" vs "randomised" manipulation is because it might prove to interact with the "same" vs "unique" manipulation: in the "blocked"/"same" set of conditions, I might expect the predictability of the noise to be at a maximum, each of the two conditions contributing to it in its own different way. On the other hand, in the "opposite" set of conditions ("randomised"/"unique"), I expect the noise to have minimum predictability and thus performance to be low. Alternatively, it might be that the "same" condition – because it "tells" the observer everything there is to know about the noise – incorporates, in a way, the predictable notch width that the "blocked" condition offers.

6. MATHEMATICAL SPECIFICATION OF STIMULI

6.1. Visual stimuli

6.1.1 The target stimulus

The target stimulus was a 512 pixels wide x 512 pixels high sinusoidal *grating* multiplied by a two-dimensional *Hanning window*. The output of this operation is presented in Figure 11 below.

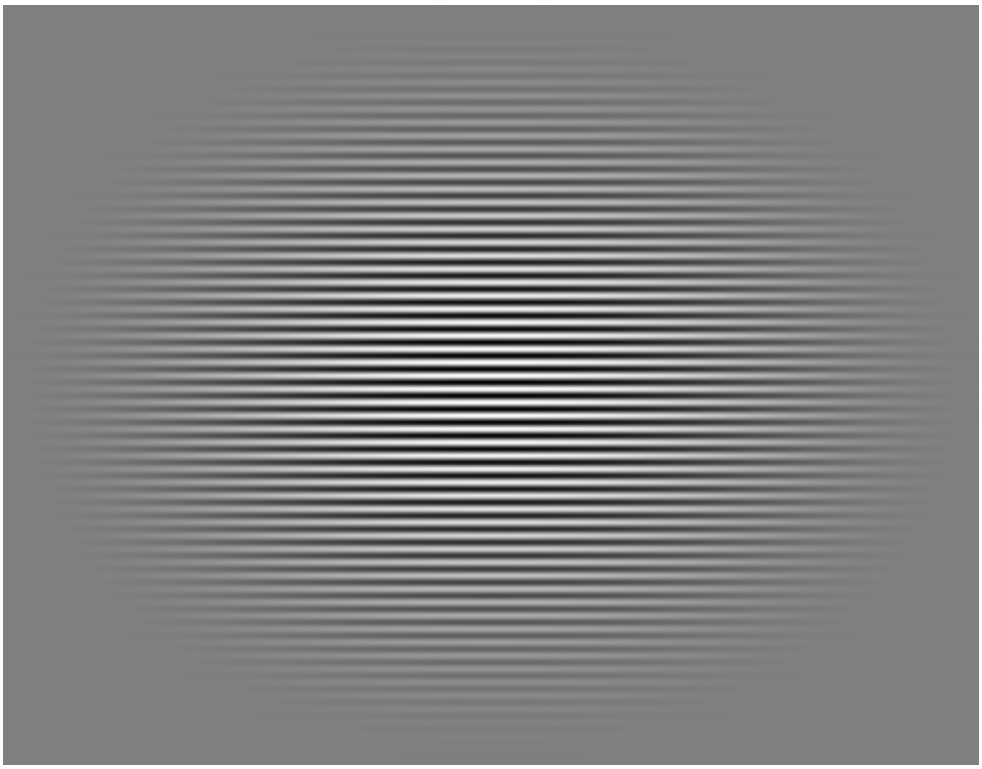

Figure 11: Target stimulus.

The next two sections describe in detail the two factors that are multiplied to obtain the target: the *grating* and the *Hanning window*.

6.1.1.1 The grating

My experiment could have been run with a *vertical* as well as with a *horizontal* grating. However, the reason why I chose the stripes in the grating to be *horizontal* and not *vertical*, is because the electron beam – that illuminates phosphors on the inside of the tube and thus creates the image – has a *horizontal* trajectory, and thus, when creating an image with *horizontal* stripes, the luminance remains constant for any one given line; otherwise, having *vertical* stripes would have meant that the luminance had to change several times per trajectory line, which would have been likely to lead to dependencies between adjacent pixels, and thus to a non-linear representation of the spatial stimulus.

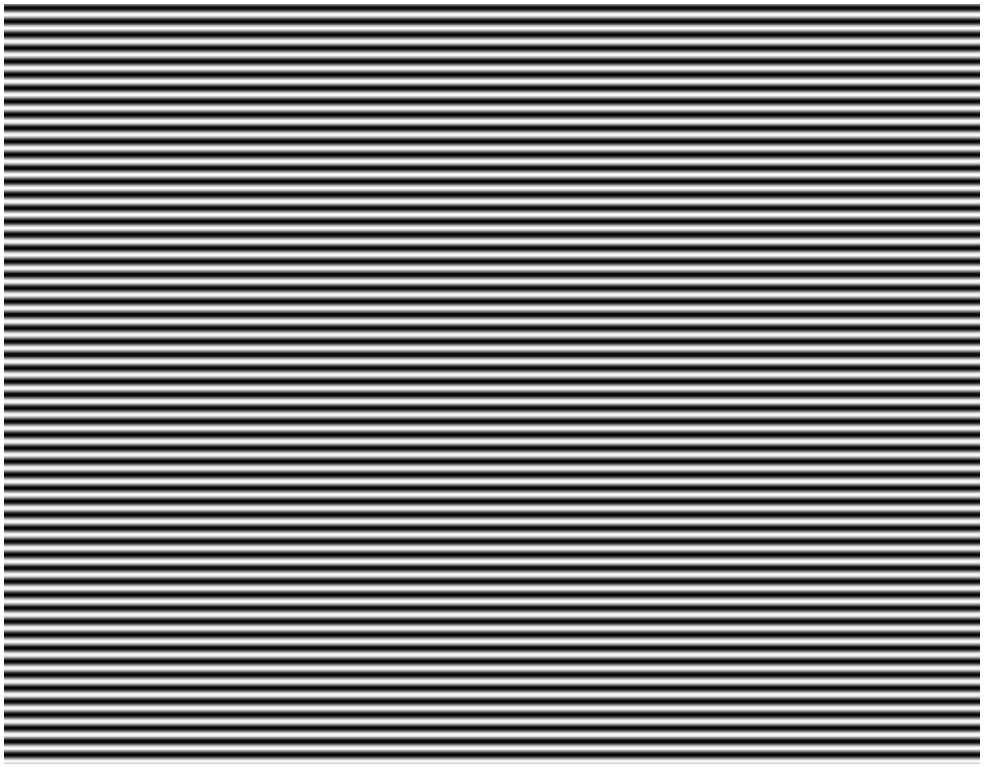

Figure 12: Sinusoidal grating with a spatial-frequency of 57 cycles per image. This (windowed) grating represents the stimulus used in the experiment; when presented on the monitor with a width/height of 25 cm and viewed from a distance of 1m (as was the case), the spatial frequency can be expressed as 4cpd.

I chose for this grating a spatial-frequency of *4 cpd*, because the maximum sensitivity (or, equivalently: minimum required threshold) of the visual system is achieved at a spatial-frequency just below this value (see section 3.2 on the contrast sensitivity function).

The reason why I chose to round *up* to 4 cpd is because the ideal stimulus would be one which has an infinite number of cycles, and therefore not displayable on a monitor. Also, a small value of spatial-frequency for the target would have lead to a lower limit of the notch being very close to zero on the frequency axis, which would have made the notch very difficult to see in a plot of the noise.

Given the viewing distance of 1 metre (see more in section 7.2.3.1), the spatial-frequency of the target can be expressed, equivalently, as 57 cpi (cycles per image). This value will visibly appear in later sections, in plots of spectral distribution of noise centred around the target.

6.1.1.2 The Hanning window

In order for the Fourier transform of a visual signal to be completely determined mathematically, the signal would have to be of an infinite length. Real-world stimuli are, however, of finite length, which can be seen as the result of multiplying the ideal, infinite-length signal with a hard-edged window that only lets a portion of it through. As a result of this, unwanted spectral components may appear, which can lead to artefacts in the spectrum of the (finite-length) signal such as *ringings* and *aliasing*. On the other hand, if the ideal, infinite-length visual signal is multiplied with a *smooth*, bell-shaped window function (such as a *Gaussian window* or a *Hanning window*), then most of these unwanted components are eliminated, thus leaving a "cleaner" Fourier spectrum of the finite-length signal.

I will use a *Hanning* window as a multiplication factor for the two visual signals I will be using (the sinusoidal grating and the noise masker). The Hanning window is described by the function:

$$w(R) = 0.5(1 + \cos\frac{2\pi R}{N}),$$

where N is the number of pixels across any dimension of the image (X or Y) and R is a position vector in the X-Y plane:

$$R = \sqrt{X^2 + Y^2}\ .$$

This window function produces values in the range [0; 1], and if those values are made to represent luminance, the following visualisation of the Hanning window is obtained (see Figure 13):

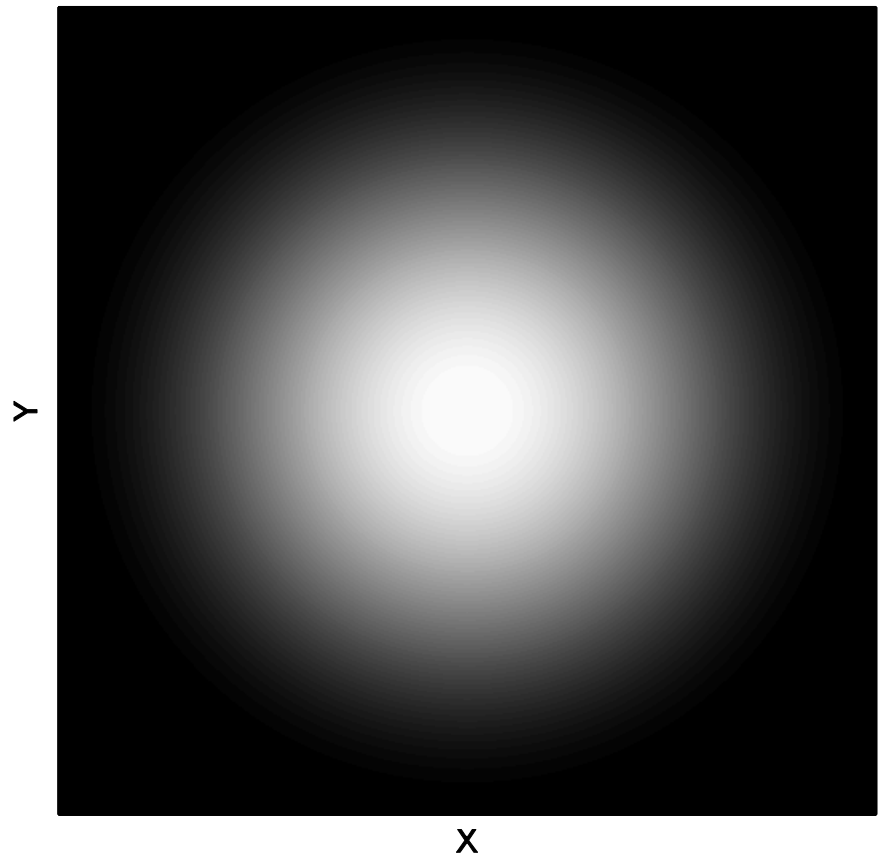

Figure 13: Two-dimensional representation of a Hanning window. White corresponds to maximum values (i.e. 1) and black corresponds to minimum values (i.e. 0) with intermediate values represented by a linear grey-scale ramp.

The effect that the multiplication by a Hanning window has on the spectrum of the signal has been described above. As a consequence of the change in spectrum, there is also an effect on the appearance of the image. Namely, the Hanning window acts as an *envelope* of contrast, leaving the central parts of the signal unchanged, and gradually and radially fading its contrast towards the peripheral parts. This is shown as a diagram in Figure 14 below, which summarises how the target is formed:

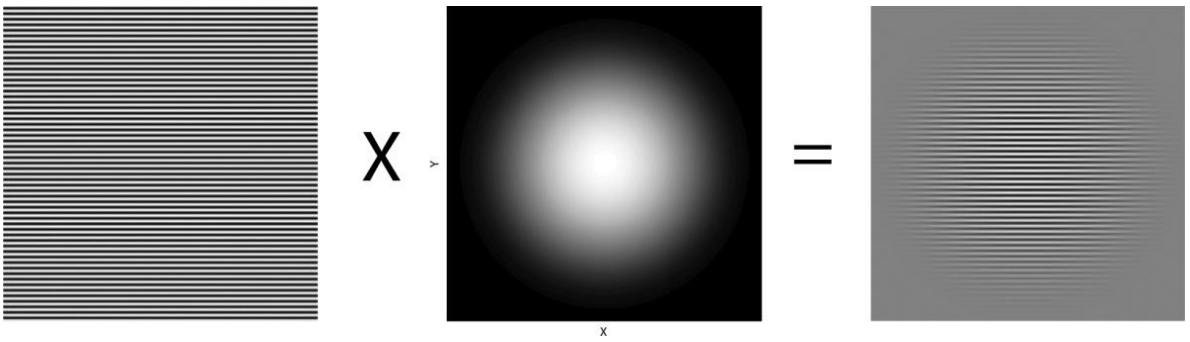

Figure 14: Multiplication of a visual signal (image) by a Hanning window eliminates most unwanted spectral components and has the visible effect of blurring the edges of the image.

In the following sections, I will use the term "Gabor patch" as a shorthand for this type of stimulus that I use here as a target, even though formally, a Gabor patch is really obtained by multiplying the sinusoidal grating with a *Gaussian* window rather than with a *Hanning* window. The reason why I chose the Hanning window over the Gaussian window

is to maintain consistency with prior literature in which similar stimuli were used (Henning & Wichmann, 2007; Losada & Mullen, 1995; Smithson, et al., 2009, under review).

6.1.2 The masker

The masker was in the form of visual noise, generated in the spatial-frequency domain using a random number generator with values drawn from a normal distribution with mean 0 and standard deviation 1.

Each sample of noise took the mathematical form of a complex number[23]. The random number generator was used to create two vectors of random real numbers, which defined the real and the imaginary parts respectively of each sample of noise. Applying the inverse Fast Fourier Transform (iFFT) to the complex noise produced a time series of N[24] numbers which, if plotted, has a chaotic ("noisy") appearance:

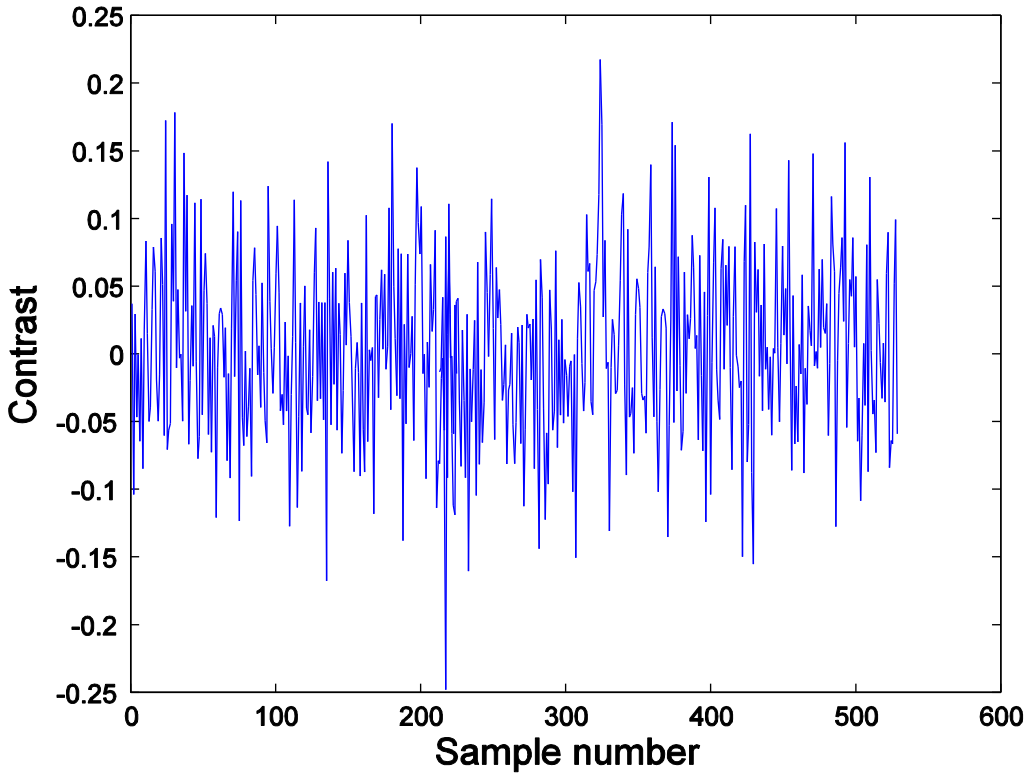

Figure 15: Plot of a randomly generated time series of noise (512 samples), with no notch in it.

The two random vectors met certain conditions[25] which ensured that the time series was real, i.e. consisted of numbers which did not have an imaginary part.

[23] I will remind the reader the definition of a complex number, first given in footnote 6 (section 3.1.1): A complex number is one that can be written in the form **a** + **bi**, where **i** is called the *imaginary unit* (defined by $\mathbf{i}^2 = -1$) and where **a** and **b** are real numbers called the *real part* and the *imaginary part* of the complex number, respectively.

[24] N is the number of pixels that make up one visual stimulus. Since the Fast Fourier Transform (which I used when creating the stimuli) works better with numbers which are powers of 2, I chose N to be **512**, which also makes for an image which fits on a monitor using the 800x600 resolution.

[25] The imaginary parts vector had its first and last elements equal to zero, and the value at –f was the complex conjugate of the value at +f.

I *scaled* the time series using a *scaling factor* chosen such that most values in the time series would fall between -1 and 1, with only a small percentage of them (around 0.1%) exceeding this range and having to be *clipped*. If too many values needed clipping (because of a large scaling factor), this would "fill in" the *notch* that needs to be introduced in the noise, i.e. make it shallower. On the other hand, if too small a scaling factor is chosen, the *noise* will be too weak.

Iterative calculations have led me to a scale factor value which best satisfied the above conditions. Thus, a plot of a typical *scaled* time series would look like this:

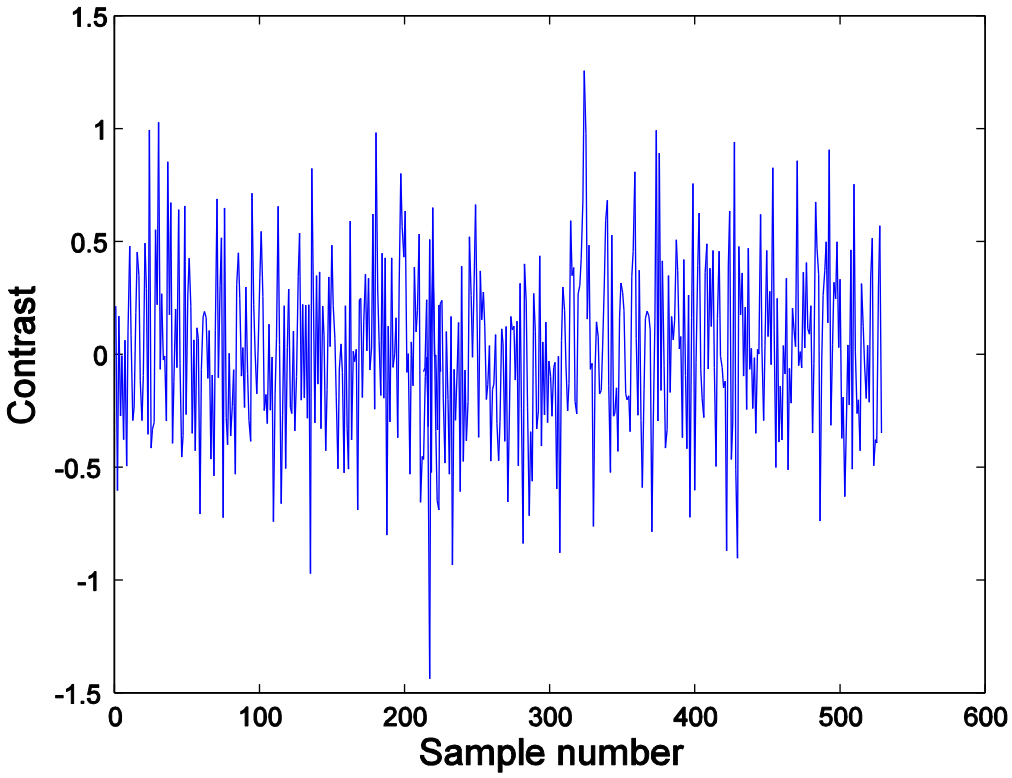

Figure 16: Plot of a randomly generated time series of noise (512 samples), with no notch in it, after having been scaled. Most values are within the interval [-1; 1], and are centred around 0 because the Gaussian distribution used to generate the noise samples had a mean of 0.

Before being displayed on the monitor, the noise had to undergo three mathematical operations:

1) *multiplication by a Hanning window*, to ensure "softness at the edges", as was done with the Gabor target. The previous section (6.1.1.2) explains why this multiplication is important;
2) *clipping*, which, as mentioned above, meant rounding the values which were below -1 to -1, and those that were above 1 to 1;
3) (a second) *scaling* followed by *rounding*, which transformed the real values in [-1; 1] to integer values in [0; 255]. This operation was needed because the function from the CRS Toolbox that displays a numeric matrix on the screen requires the matrix's values to be integers between 0 and 255.

In order for the noise thus generated to be useful as a masker, it needs to have a notch introduced in it, whose width can then be varied as an experimental manipulation.

The stopband of the noise (i.e. the *notch*) should be centred around the spatial-frequency of the Gabor target, while the passband (i.e. the frequencies which are outside of the notch) should be of a relatively constant level.

In order to check that the noise that I was going to display on the screen indeed met these conditions, I have performed – successively – the three operations described above upon 100 exemplars of noise (i.e. time series) and then averaged their Fourier Transforms (i.e. spectra), thus obtaining a more representative, averaged *power spectrum* of the noise. A plot of this spectrum for an inserted notch width of 2.0 octaves is given in Figure 17 below, using double logarithmic scales:

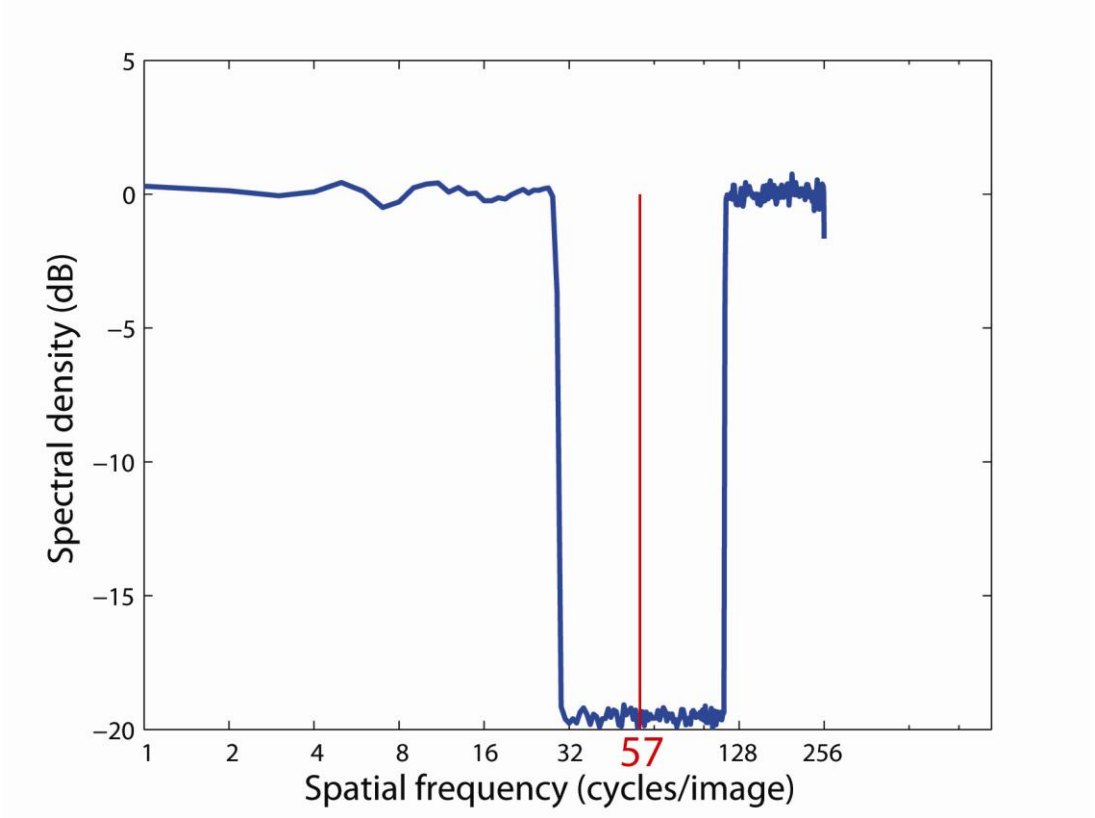

Figure 17: Shape of the frequency domain characteristic (power spectrum), averaged across 100 exemplars of noise, all having a 2.0 octaves notch width inserted around the target's spatial-frequency of 57 cpi (emphasised with a red vertical line).
One line (time series) of noise consists of 512 samples, which is why, according to the Nyquist-Shannon theorem, there are spatial-frequency components up to 256 cpi.

What has been described so far represents *one* time series of noise, i.e. 512 different values of contrast. In order to obtain the 2-dimensional visual stimulus corresponding to the noise, one such time series of noise is represented as a vertical line, and that line is repeated horizontally 512 times, to obtain a 512 x 512 pixels visual stimulus of noise, which looks like this (see Figure 18):

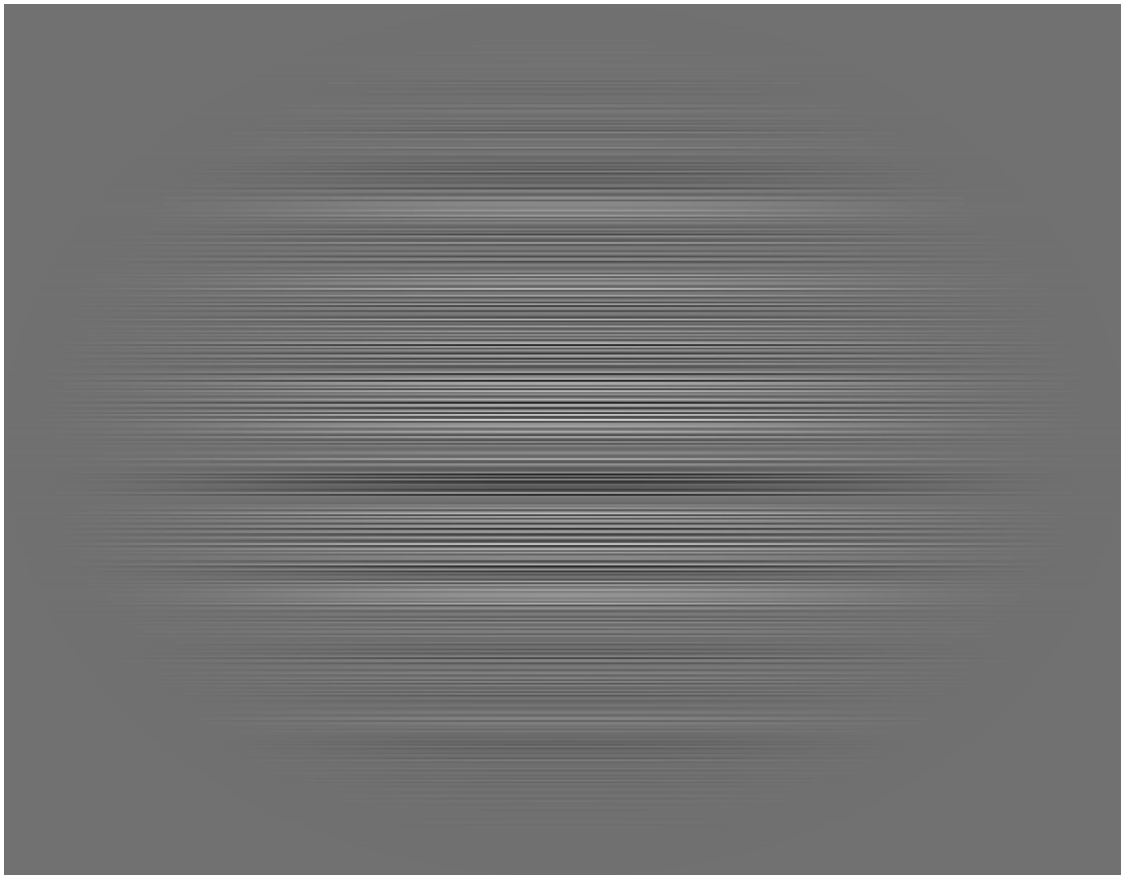

Figure 18: The visual noise stimulus, as it appeared on the screen. The noise has been multiplied by a Hanning window.

As explained in the previous paragraphs, the target and the noise have each been multiplied by a Hanning window. The reason why this is not a problem (and does not lead to an undesirable second multiplication) is that multiplication is a *distributive* operation, such that the adding of the target multiplied by a Hanning window and the noise multiplied by a Hanning window can be equated to the target added to the noise and both of them being then multiplied by the Hanning window.

6.2. Auditory stimuli

6.2.1 The target stimulus

The target stimulus in the hearing experiment was constructed in a way that was mathematically analogous to the one in the vision experiment, only with different numerical characteristics. Namely, the target stimulus was a 250 ms-long, 2 kHz purely-sinusoidal tone that was passed through a Hanning window.

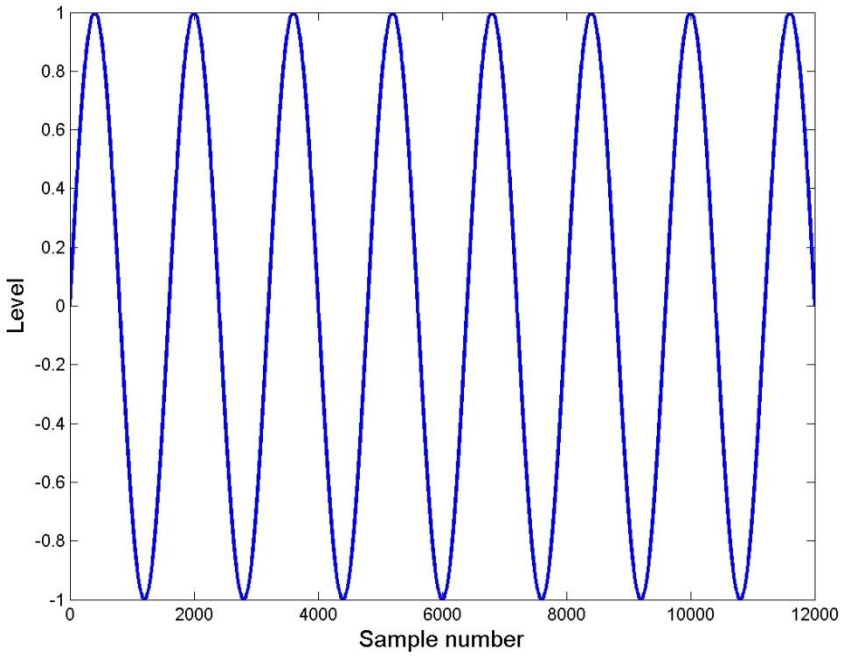

Figure 19: A qualitative representation of the pure (unwindowed) sinusoidal tone. This is not representative of the actual *frequency* that was used (*2 kHz*), which cannot easily be represented in print. Rather, the graph represents a tone of the same duration (250 ms) but of *30 Hz* frequency. At a sample rate of 48 kHz, this means a total of 12,000 samples.

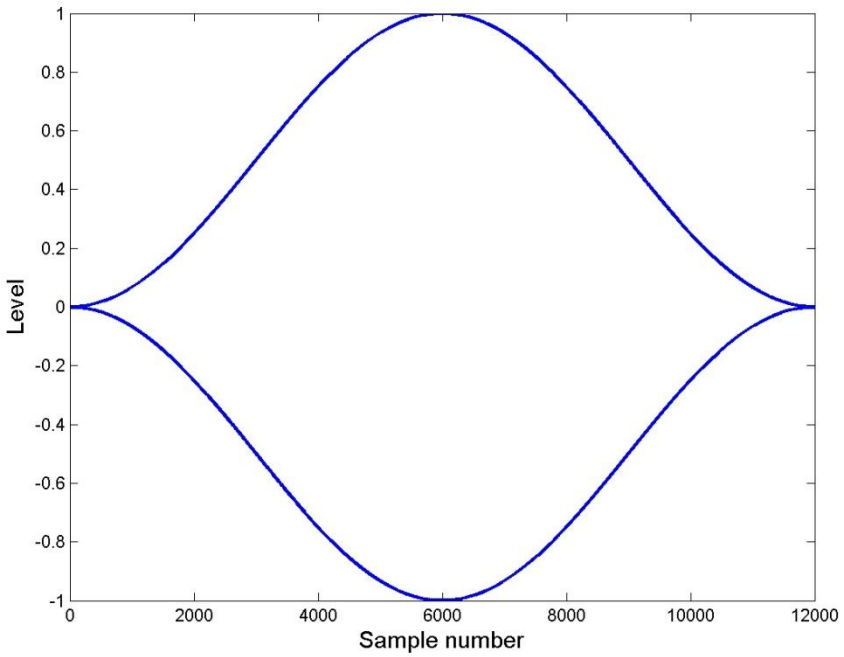

Figure 20: The Hanning window, represented at the same time scale as the tone.

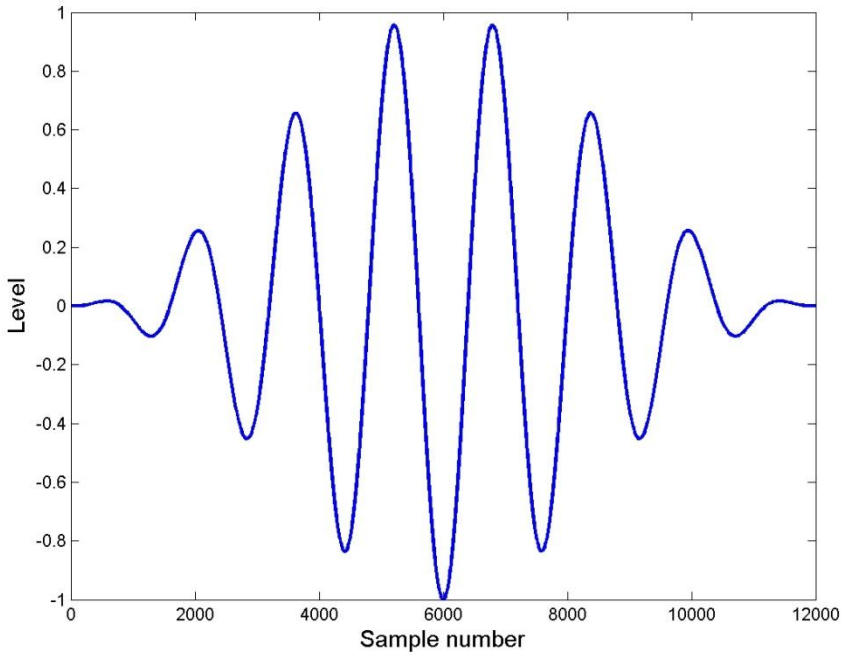

Figure 21: The windowed sinusoidal tone (i.e. after it has been multiplied by the Hanning window), represented at the same time scale as the tone and the Hanning window.

The *amplitude* (or: *level*) of the tone was varied by the staircase that controlled the experimental session, this constituting one of the main experimental manipulations.

The Hanning window was defined in analogy with the vision case (see section 6.1.1.2), only the Hanning window was one-dimensional in this case rather than two-dimensional as it was in the vision case. The multiplication has the same effect of "softening the edges", as can be seen in the diagram in Figure 22 below, which is analogous to the one in Figure 14:

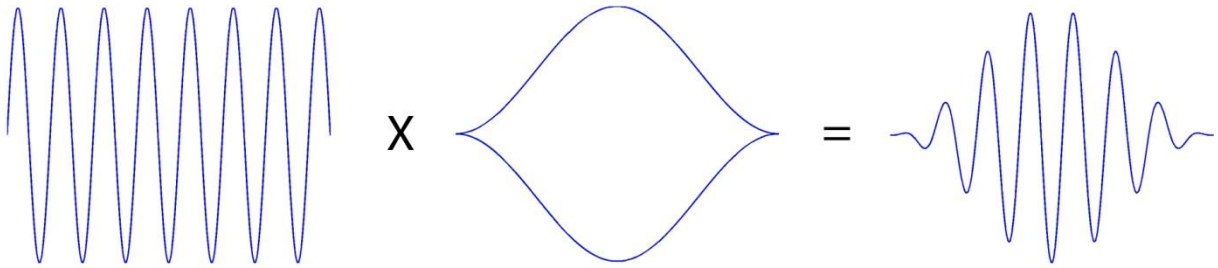

Figure 22: Multiplication of an auditory signal (tone) by a Hanning window eliminates most unwanted spectral components and has the visible effect of blurring the edges of the waveform of the tone. Again, the temporal scales for both the tone and the window are not representative of the actual stimuli used, but have been scaled to allow a more meaningful visualisation.

6.2.2 The masker

Noise was made in the same way, i.e. generated in the frequency domain etc, but was played as a time series (i.e. sound) rather than displayed as an image, as was the case with the vision experiment.

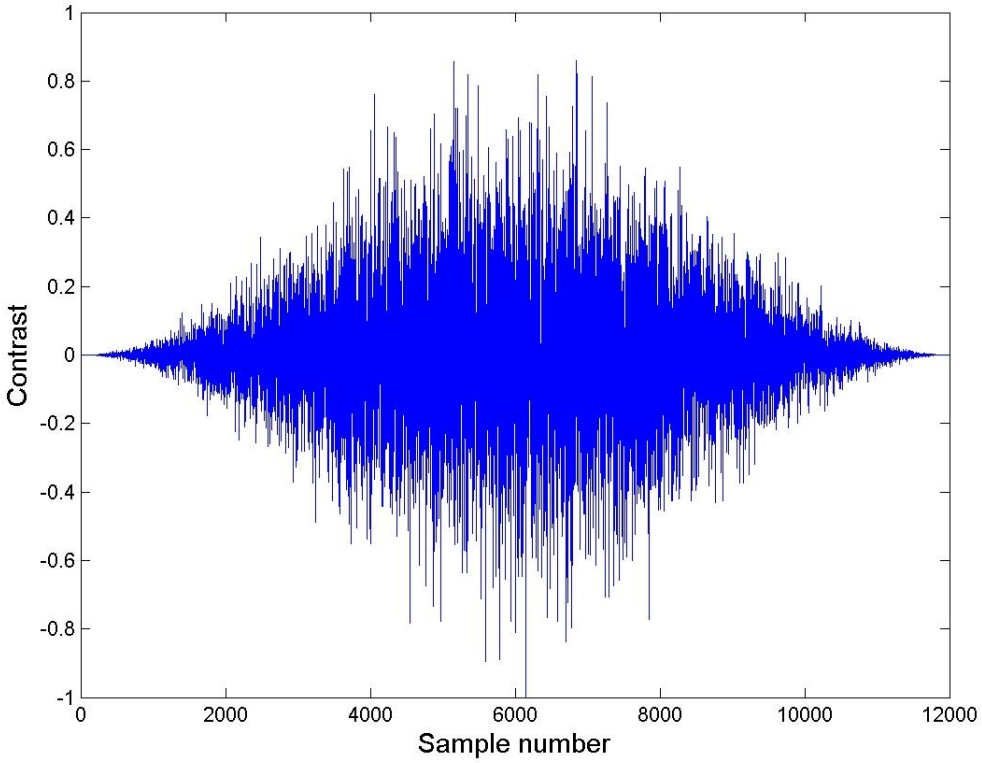

Figure 23: Plot of the noise used as a masker in the hearing experiment (with notch width = 0.0 octaves). The noise has been multiplied by the same Hanning window used to multiply the sinusoidal tone. Here the temporal scale for the noise *is* representative of that used in the experiment.

The noise was mathematically added to the target in order to create the noisy tone interval.

6.3. Analogies between vision and hearing

In order to better understand the analogies between vision and hearing that make stimuli in the two cases be so similarly defined, the table below gives the equivalents in hearing of a number of notions used in vision, as well as of a number of parameters used in the two experiments.

Table 4: Analogies between parameters of the vision experiment and those of the hearing experiment.

Vision	Hearing
image	sound (or: wave form)
contrast	amplitude (or: volume)
pixel	sample
sinusoidal grating	sinusoidal tone
spatial-frequency [cycles/degree]	temporal-frequency (or: pitch) [cycles/second]

spatial extent (size) [degrees of visual angle]	duration of auditory stimulus [seconds]
signal+noise interleaving through page cycling	signal+noise interleaving through algebraic summation of signal+noise
blank page (50% grey)	silence
sampling frequency (i.e. resolution)	sampling frequency (i.e. soundcard's sampling frequency)

7. VISION EXPERIMENT

7.1. Introduction

I start with the vision experiment, and then design a similar experiment in hearing. Finally, I will draw conclusions based on the parallels (both similarities and differences) between the two.

The experiment measures contrast detection thresholds for spatially-defined visual targets, in the presence of notched noise, and as a function of the width of that notch. I manipulate the way in which different notch widths are grouped over trials and across intervals of the same trial.

7.2. Methods

7.2.1 Observers

Three normal, non-naïve observers were used for the vision experiment, including the author (TP) and his supervisor (HS):

- Tudor Popescu (TP)
- Hannah Smithson (HS)
- Wayne Smith (WS)

The three observers each collected the same amount of data, a process which extended across a time span of several weeks to months. The time lapse between two consecutive "sittings" (i.e. block of 5 sessions) was not fixed for any given observer and was also not the same across observers.

Observers TP and WS had normal vision while observer HS had corrected-to-normal vision.

7.2.2 Task

In all stages of the vision experiment, my aim was to measure contrast thresholds. The target to be detected was a 4 cpd Hanning-windowed grating of 512 x 512 pixels extent.

The task that each observer was asked to perform on each trial was in the form of a temporal 2AFC – *two-alternative forced choice* between two intervals (or 2IFC – *two-interval forced choice*): the observer had to indicate which of the two stimuli displayed consecutively on the screen contained the target stimulus. One of the intervals contained the target embedded in notched noise, while the other interval only contained the notched noise. The order that the two intervals were presented was chosen randomly for each trial.

The duration of each of the two intervals was 50 ms, and they were separated by a 500 ms blank interval, which consisted of a uniform field of 50% grey luminance – the same luminance which was used as a background for the two intervals.

Between two consecutive trials there was another 500 ms blank, which was long enough for the after-effects of the previous image to have dissipated, without overly increasing the duration of the session.

A beep sound with the same duration as that of the intervals (50 ms) was played simultaneously with each interval.

Feedback was given to the observer after each trial, in the form of another beep, audible immediately after the observer had given the response, which had a low frequency if the answer had been incorrect and a high frequency if the answer had been correct.

I chose to give feedback because by doing so, the observers are more easily able to ascertain which of the cues they had been using to perform the task are correct, and thus to more quickly reach their plateau[26] (hopefully without influencing the absolute value of this plateau).

7.2.3 Equipment

7.2.3.1 Display equipment

The stimuli were presented on a Mitsubishi Diamond Pro 2070SB CRT monitor, used at a frame rate of 160 Hz and at a resolution of 800 x 600 pixels. The 512 pixels wide x 512 pixels high stimuli displayed on the monitor measured 25 cm in both width and height. The viewing distance was held constant at 1 m, which meant that the displayed stimuli subtended a visual angle of 14.22° x 14.22°.

The video card connected to the monitor was a *VSG (Visage)* manufactured by *Cambridge Research Systems Ltd.,* which was programmed to display stimuli by means of a *Matlab* interface (CRS Toolbox for Matlab).

7.2.3.2 Display of stimuli

As mentioned in section 7.2.2, each trial consisted of two intervals containing the same or unique samples of noise, always presented at 100% contrast, with one of the two

[26] A plateau is an asymptote of one's performance.

intervals randomly having the Gabor target added to it, which was of varying (and typically, much lower) contrast.

When displaying a certain stimulus – target+noise (i.e. noisy Gabor) or noise alone – on the monitor, the video card controlled the contrast of the stimulus by means of a lookup table (LUT). If the noisy Gabor interval were obtained by directly adding the noise to the Gabor target, then a single LUT would be used for both the 100%-contrast-noise and the variable-contrast-target, which would halve the precision with which the target's contrast could be controlled.

Therefore, to work around this problem, the interval containing the target was obtained not by adding the two components (*noise* and *Gabor target*) together in the same image, but by displaying them on alternate frames[27]. Specifically, *odd* frames contained the noise while *even* frames contained the Gabor target.

In order to maintain consistency in how the two intervals were generated, the interval *not* containing the target was also made up of alternating frames, with odd frames containing the noise and the even frames containing a blank page (uniform field) equivalent to a zero-contrast target.

I had to ensure that putting target and noise on alternate frames is equivalent to adding them up, before presenting them. In order for the target and noise to be integrated, the duration of a target/noise pair of frames needs to be smaller than the integration time of the visual system (roughly 100 ms).

The 160 Hz frame rate used for the monitor meant that two frames would last 12.5 ms, which means that the condition stated above is met by almost an order of magnitude.

7.2.3.3 Response box

The observers' responses were recorded using a wired Cedrus RB-530 response box with five buttons, of which only two were used: the observer had to press the *left* button in order to indicate that the target was in the *first* interval, or the *right* button to indicate that it was in the *second* interval.

7.2.4 Calibration

In order to be able to accurately control the parameters with which the monitor displays images, *gamma correction* needed to be done, which ensures that the luminance produced by the monitor depends *linearly* upon the voltage applied between its electrodes.

Gamma correction is done by applying several voltages and measuring (using a linearised photo-diode – ColorCAL, supplied by CRS Ltd) the resulting luminance, then fitting a curve through all these points. Since this curve was not initially linear, I had to use the gamma correction software supplied by CRS to apply a linearisation procedure,

[27] For this I have made use of the VSG's ability to *cycle* video pages and their associated LUTs.

which was implemented by modifying the voltages used by the various values in the look-up table.

Mathematically, if the initial monitor curve can be described by the relation:

$$L = V^x,$$

then the correcting curve needs to have the equation:

$$L = V^{-x},$$

such that, when implemented into the monitor (an operation mathematically equivalent to multiplication), the resulting curve would be a line, described by the relation:

$$L = V.$$

After this has been done, the luminance provided by the monitor would depend linearly upon the applied voltage, which meant that the monitor has been gamma-corrected.

7.2.5 Procedure

The trials were displayed using a staircase procedure, with five interleaved staircases running simultaneously and independently of each other. The decision which of the five staircases displayed a trial was made at random, with values drawn from a uniform distribution, i.e. each staircase was equally likely to be chosen.

For one given staircase, the contrast of each trial was determined using a method called "accelerated stochastic approximation" (Treutwein, 1995), whereby the contrast of the current trial, X_{n+1}, is related to the contrast of the previous trial, X_n, by the following formula:

$$X_{n+1} = X_n - \frac{c}{2 + m_{shift}}(Z_n - \phi), n > 2$$

where

- **c** is a constant that has the value 0.5,
- $\mathbf{m_{shift}}$ is the number of reversals[28] that have occurred so far,
- $\mathbf{Z_n}$ is the correctness of the previous response (1 for correct and 0 for wrong), and
- **Φ** is the target probability, i.e. the *performance level* (see section 2.2.1 for definition), which I took to be 0.75 (or 75%).

The above formula requires an initial contrast (X_1), and I have chosen this to be 40% for all staircases, as I have found that this value minimises the running time of a staircase for my specific task.

A staircase finished when it had reached its sixth reversal. However, after this point, it would still continue to display trials, in order to not reduce the unpredictability of the next trial for the observer. All trials ran after the sixth trial were ignored, and the

[28] A reversal is a trial for which the response of the observer has been correct, where the previous response had been wrong, or vice versa.

threshold was calculated by averaging the last four reversals, i.e. the third, fourth, fifth and sixth reversals.

After each trial, some of the more relevant data collected for that trial, pertaining either to the staircase that the trial belonged to, or to the trial itself, were written to a CSV[29] file. This allowed for subsequent analysis of the data acquired in that session, and also made it easy to plot any data that I deemed relevant.

7.2.6 Conditions

My experiment has three independent variables, and I measure thresholds as a function of the conditions defined by combinations of these variables. The three variables are:

1. *Notch width* of the noise, and I manipulate this parameter in order to measure "channel" bandwidth. The five possible values of notch width were: 0 octaves, 0.5 octaves, 1 octave, 1.5 octaves and 2 octaves. These values are equidistant on a logarithmic scale and they provide a good enough resolution on the graph that plots (on the vertical axis) the threshold values obtained for each of those notch widths (on the horizontal axis) so that a curve can be fitted with a good enough approximation through those five points;
2. A parameter that I refer to as *"blocked" vs "randomised"*, which indicates whether thresholds for a particular notch width were obtained in pure sessions defining a particular noise context ("blocked"), or whether noise properties varied from trial to trial ("randomised"). This manipulation allows me to assess whether noise properties can be learned under conditions in which noise context is held constant;
3. A parameter that I refer to as *"same" vs "unique",* which indicates whether a fresh exemplar of noise is used in each of the two intervals of the trial ("unique"), or whether the exemplar is repeated for the two intervals ("same"). In this manipulation, I ask whether performance is improved if the observer is given the opportunity to predict the noise in the second interval, because it is identical to the one in the first interval rather than simply sharing the same definition in terms of notch width.

All three observers were well aware of the existance of the two conditions, however before a given session, they would not be made aware of the conditions under which that sessions would run. This means, for example, that they would not know whether the noise in the two intervals of a trial is going to be the same or different.

The staircase controlled the contrast of the target on a trial-by-trial basis, but the properties of the noise were fixed for a given staircase. Specifically, the notch width of the noise was, in the blocked conditions, the same for all five staircases, or was different for each staircase in the randomised condition.

I defined one session to be a complete run of all five interleaved staircases that measure thresholds for their respective notch widths – the way each staircase of the session was assigned a notch width (either in a "blocked" or in a "randomised" condition) gave name to the session.

[29] Comma-separated values, a format readable by spreadsheet software.

The experiment was performed by acquiring data from each of the three observers, in the form of one block of five sessions (a "sitting"), which meant an average of 40 minutes per sitting per observer.

Since there were five possible notch width values, each sitting had to contain five sessions, so that all five notch width values could be represented (for a "blocked" sitting); the same number of sessions was also used for a "randomised" sitting, in order for the two conditions to be equally represented and thus eligible for further comparison.

The *noise condition* ("same" or "unique") was alternated at every other sitting, while the *presentation condition* ("blocked" or "randomised") was alternated at every sitting. So for example, an "SB" sitting would be followed by an "SR" sitting, then by a "UR" sitting, then by a "UB" sitting and so on. As this may seem difficult to visualise without a clear schematic representation, a full outline of the order of sessions has been included in Appendix A.

For the "blocked" sittings, the order in which the notch width values would be distributed to the five sessions was changed at every other "blocked" sitting, according to a Latin square design[30]. So a first pair of "B" sittings (for example, an "SB" sitting and, later on, a "UB" sitting) would run their five sessions with notch widths of 0.0, 0.5, 1.0, 1.5 and 2.0 octaves, then the next pair of "SB", "UB" sittings would use 0.5, 1.0, 1.5, 2.0 and 0.0 octaves, and so on. See Appendix A for more details.

Finally, within a sitting, the order of the "blocked" and "randomised" sittings was alternated according to an A-B-B-A repetition pattern. So for example an "S*B*" sitting followed by an "S*R*" sitting would then be followed, in order, by a "U*R*" sitting and a "U*B*" sitting. Again, to gain a clearer overall picture of this, see Appendix A.

Also see Appendix B for a note on the counter-balancing of trials throughout the experiment.

7.3. Results

Throughout this Results section, I will plot thresholds using the *logarithm* of the contrast values rather than the contrast values themselves. The reason why I do this is because human perception (e.g. vision and hearing) works logarithmically rather than linearly. This becomes evident in the form of the relationships that can be determined between perception and stimulus strength, such as Weber's law, which states that perception is a logarithmic function of stimulus strength.

Consequently, I will refer to these values as "*log contrasts*" or "*log thresholds*". The curves therefore represent log contrast, however when labelling the y-axes, I change the log contrast values back to plain (non-logarithmised) values of contrast, so that the tick marks on the y-axes can be more easily interpreted as being 0-to-1 contrast values rather than the more abstract log contrast values.

I should remind the reader that I will often use shorthand to refer to both experimental conditions, with a slash separating the two possible levels of the condition:

[30] A Latin square design (or: rotation experiment) is a pattern in which each of n levels of a factor (variable) is represented once in each column and once in each row of a square n x n matrix (D. Campbell, Stanley, Gampbell, & Stanley, 1969)

S/U for the noise condition, which can be either **S** for "same" or **U** for "unique"; **B/R** for the presentation condition, which can be either **B** for "blocked" or **R** for "randomised". Any of the four possible combinations of those conditions will also be referred to in shorthand, with the two levels written one after the other with no separator (e.g. **SB** means noise condition "same" *and* presentation condition "blocked"). Finally, I will be referring to *notch width* as **nw** and to *estimate number* as **en**, when it appears as a factor in an ANOVA.

What I essentially did in this experiment was to measure the contrast required to detect a target stimulus embedded in noise, as a function of the notch width in that noise. Using the data I have collected from my three observers under the various conditions, I am going to present my main findings with regards to the vision experiment by producing plots that can visually suggest particular trends, and then by doing the relevant statistics on the data to see whether the apparent trends are indeed genuine.

I will divide the presentation of the results into sections that first look at the extent to which my main hypotheses were supported by the data, and then look at whether and how the performance of the observers evolved as they gained experience with the task and potentially exhibited perceptual learning.

7.3.1 Main analysis

In this section I will first visually explore the data by plotting it in various ways, then I will I use statistics – done on the entirety of my data as well as on each individual observer – to formally test my initial hypotheses. I shortly remind the reader that this consisted of several research questions:

- whether there would be a main effect of the B as opposed to the R condition, and whether B/R would interact with nw;
- whether there would be a main effect of the S as opposed to the U condition, and whether S/U would interact with nw;
- whether the B/R factor would be most apparent in the U condition, because presumably under the S condition, there is no "room" for further learning.

7.3.1.1 Graphical comparison between "blocked" and "randomised"

The most straightforward way to compare the B and R conditions is to plot on the same graph the curves representing threshold as a function of notch width obtained for each of them. Figure 24 below does this for each of the three observers, separately for the S and U conditions. Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.

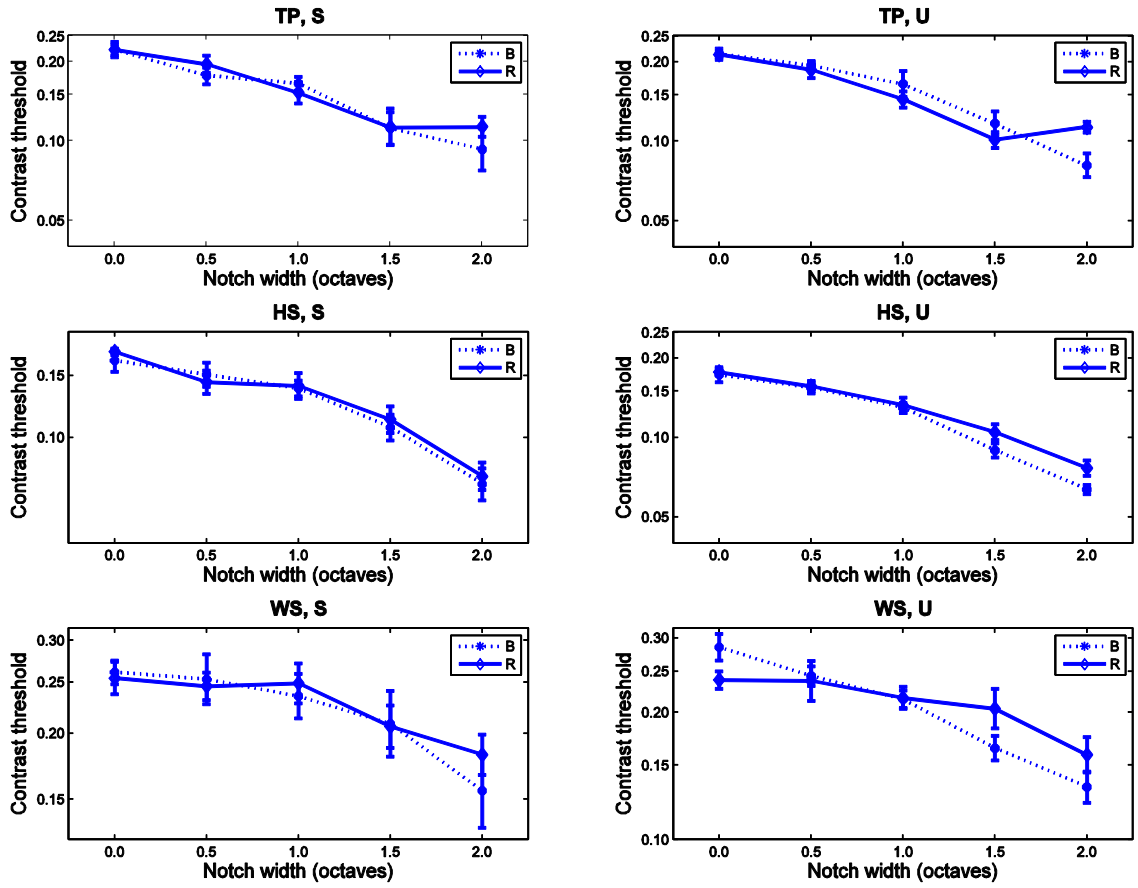

Figure 24: Graphs comparing results for the B and R conditions. The six graphs are for each of the three observers and under each S/U condition. Each data point represents the average of the five different estimates that were taken for that observer, under that set of conditions, and for that value of notch width. Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.

My hypothesis was that the "blocked" curve would lie below the "randomised" curve, indicating an ability to noise characteristic when presenting trials randomised. At first sight however, there is no such tendency that is obvious and unanimous for the graphs displayed. Observer HS comes closest, with most of her "blocked" data points lying below the "randomised" data point for the same notch width. However, the extent to these differences are significant will be studied statistically later, in the global analyses performed in section 7.3.1.4.

7.3.1.2 Graphical comparison between "same" and "unique"

An overall picture of the difference between the S and U noise conditions can be gathered by plotting thresholds (averaged across all estimates) as a function of notch width, for two different curves that represent the S and U conditions. There will be six such graphs, for the three observers and two presentation conditions (B and R), and they are shown in Figure 25 below. Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.

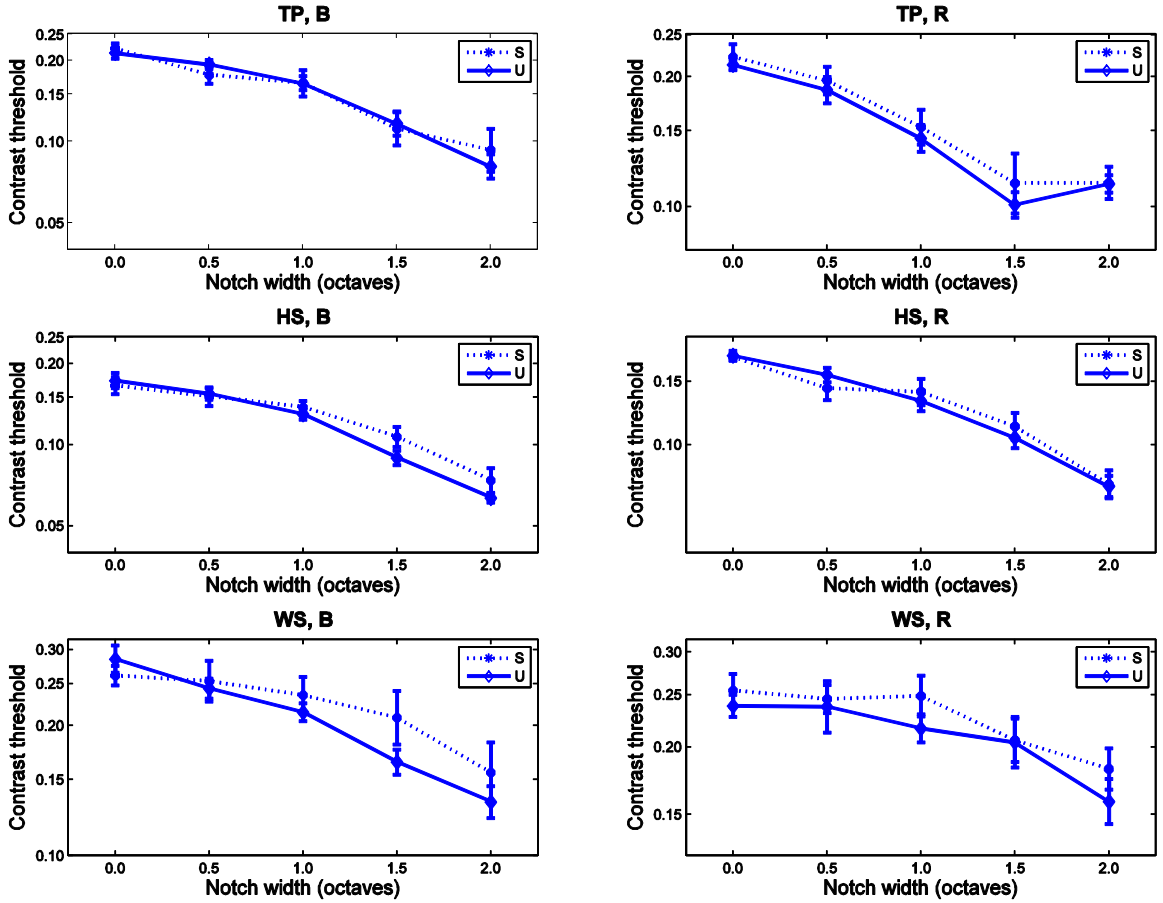

Figure 25: Graphs comparing results for the S and U conditions. The six graphs are for each of the three observers and under each B/R condition. Each data point represents the average of the five different estimates that were taken for that observer, under that set of conditions, and for that value of notch width. Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.

The phenomenon which has traditionally been used to infer channel bandwidths – namely, that *thresholds* get lower as *notch widths* increase – can clearly be seen from Figure 25 above. There do not seem to be large differences between the S and U conditions, but this will be assessed statistically in the following sections.

7.3.1.3 Individual differences manifested in the observers' performance

I was not particularly interested in investigating the individual differences between my observers, first of all because this was not part of my research question and second of all because the low number of observers that I used implied that any analysis done on this basis would probably be inconclusive. However, a quick glance at the comparative plots that can be drawn allows one to see that the difference in performance between my three observers was relatively constant under the various conditions and for the various values of notch width, with observer HS constantly obtaining the lowest thresholds, followed by observer TP and with observer WS obtaining the highest thresholds. The four plots, each comparing the performance of the three subjects under each possible set of conditions, are shown in Figure 26 below. Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.

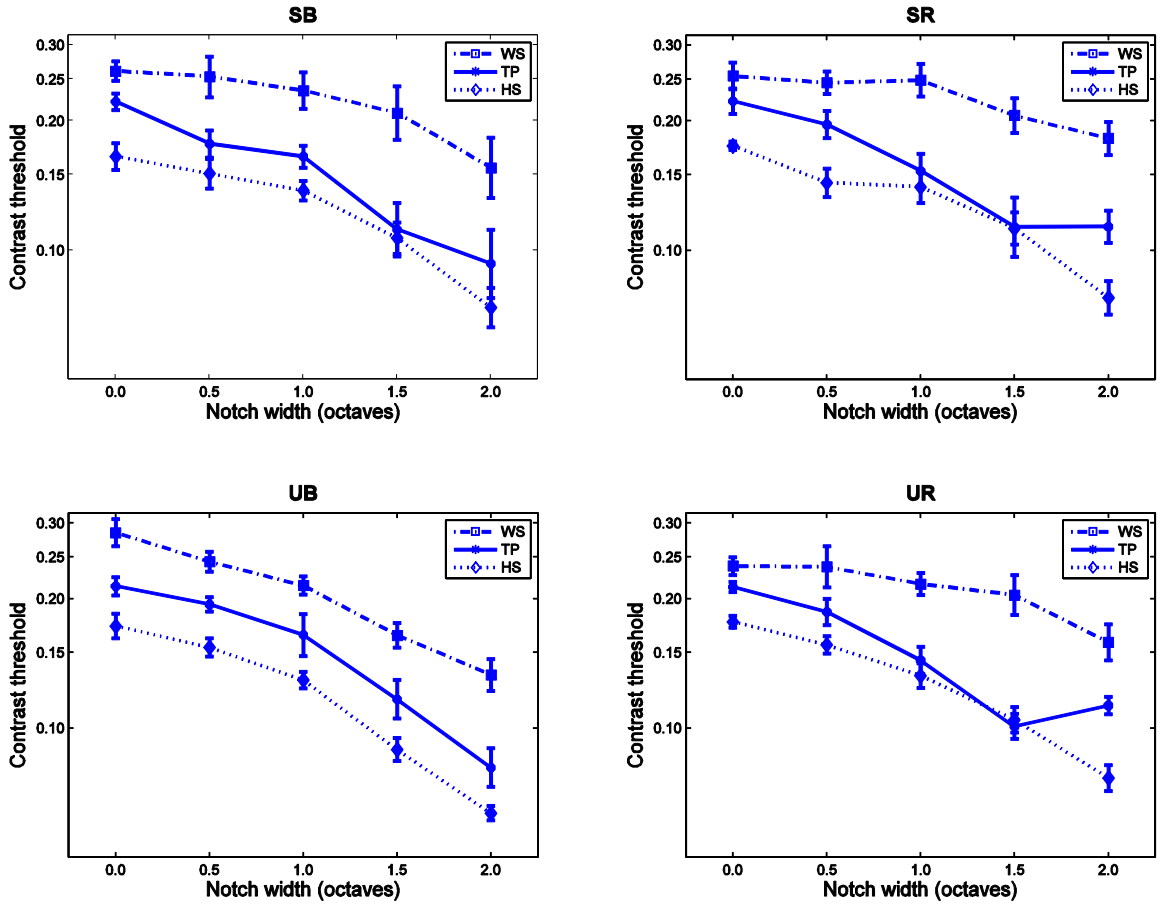

Figure 26: Plots comparing the performances of the three observers. Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.

7.3.1.4 Global analyses

I ran a 3-way within-observer ANOVA with the factors S vs U (2 levels), B vs R (2 levels) and notch width (5 levels). This was the main (global) analysis of variance done for the vision experiment, since it analyses data for all observers, and tests my initial hypotheses.

The results of this ANOVA revealed the following effects:

- There was a significant main effect of notch width, $F(4, 8) = 29.60$, $p<.001$, which suggests that log threshold values were significantly influenced by the notch width of the noise, as one would expect given the current literature. To get a feel of how large this influence is, in going from a 0.0 octaves notch to a 2.0 octaves notch gives a mean improvement in performance of .1094 in plain contrast units (or .3122 in log contrast units). Figure 27 below shows the overall dependence of threshold on notch width, after data from all observers and conditions have been averaged. Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.
  The improvement in performance with increasing notch width is confirmed by a linear trend analysis ($F(1, 2) = 39.37$, $p<0.05$).

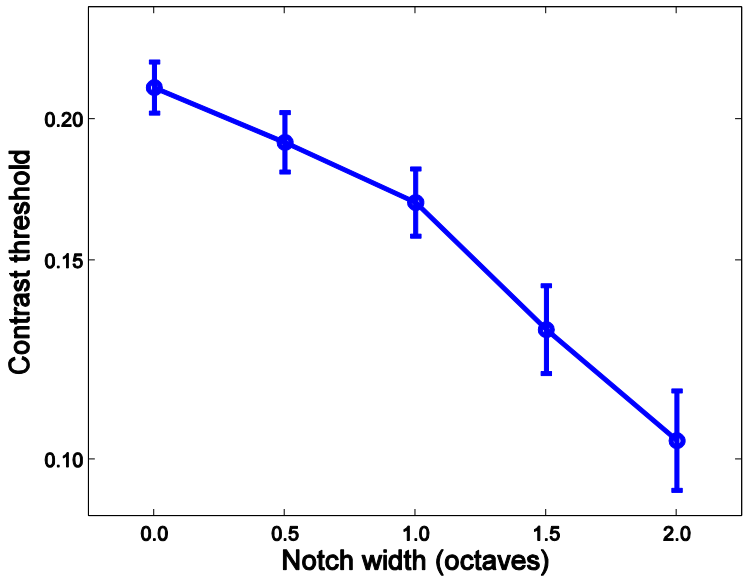

Figure 27: Overall dependence of threshold on notch width. Data points are averaged across all three observers, both noise conditions and both presentation conditions. Trend is clearly decreasing, as one would expect given the current literature. Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.

- There was a significant main effect of B vs R, $F(1, 2) = 19.15$, $p<.05$, which suggests that log threshold values were significantly influenced by the B vs R condition.
  By looking at the means, I can see that it is the B condition that produced better performance. Namely, the average log threshold value for the B condition (average taken for all observers, all notch widths and for both the S and the U condition) was -.814 and was -.798 for the R condition, giving a difference of .016 in log contrast units; in plain contrast values (non-logarithmised), this translates to a difference of 0.0058, with the mean being .1534 for B and 0.1592 for R.
- There was no significant main effect of the S vs U condition, which suggests that log threshold values were not significantly influenced by this condition.
- There was a significant interaction effect between the S vs U condition and notch width, $F(4, 8) = 6.05$, $p<.05$, which suggests that the influence of notch width on log threshold values is different in the S condition compared to the U condition.
  The fact that this interaction is significant tells me that the two curves representing S and U have different shapes (i.e. they are not parallel over the entire range of notch widths), as can be confirmed from the graph in Figure 28 below, which compares S and U across all available data. Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.

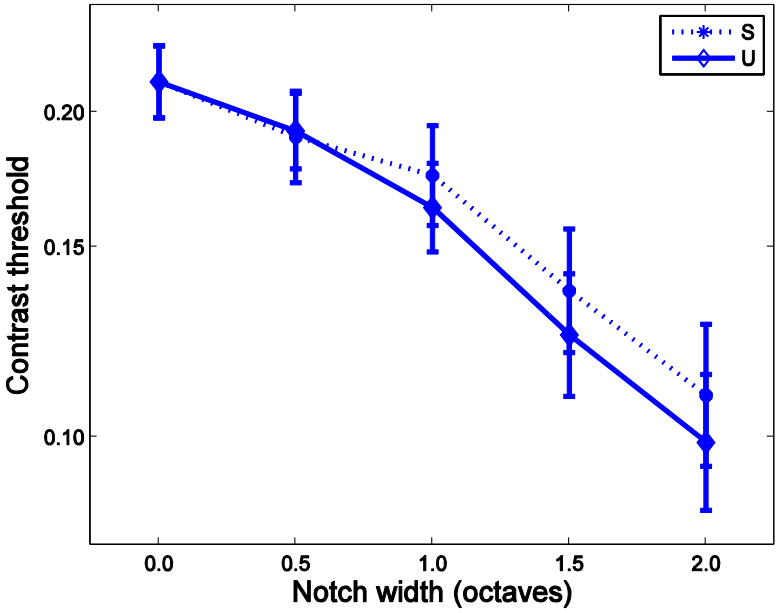

Figure 28: Graph comparing performance under the S and U conditions; all data points are averaged across all estimates, across all three observers (TP, HS, WS) and across both presentation conditions (B and R). Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.

- There were no significant 2-way interaction effects between *S vs U* and *B vs R* and between *B vs R* and *notch width*; there was also no significant 3-way interaction between *S vs U*, *B vs R* and *notch width*.

It has to be noted that, in order for the groupwise parametric statistical analyses that have been performed to yield meaningful results, it is necessary to be able to assume that the data being analysed are drawn from a *normally* distributed population. Given that there are only three observers, this assumption is rather hard to justify. Although the relevant diagnostic tests do produce a result, their power is low enough that their results might not be considered reliable. On the other hand, the analyses themselves are also very weak in power, and these two shortcomings combined mean that any positive results should be indicative of *real* effects. Also, should the assumption of normality discussed above be violated, the effect that this has on the validity of analyses is small because the dependent variables are *independently* sampled (see (Keppel & Wickens, 1982)).

7.3.1.5 Analyses per observer

After visually exploring the data using plots, I would have expected there to be a significant interaction effect between notch width and B vs R, which would have confirmed one of my research hypotheses. As shown in the global analysis however, this interaction was not significant. Still, since the experiment produced sufficient data for each observer, I will do analyses of variance for each observer individually, that will show whether there are significant effects for any one observer.

In doing analyses of variance, a factor is described as either being *within*- or *between*-observers; a *within*-observer factor is collected from repeated observations from the same observer; a *between*-observers factor describes data obtained across

different observers. The reason for describing factors like this is that the data that is collected on the same observers (i.e. *within*-observer data) is very likely to be highly related (because they come from the same observers), and it is important in ANOVA to take into account the different levels of "relatedness" for *between*- or *within*-observer factors. Here, the terminology gets confusing, because I am only doing analyses per observer, but still this issue of "relatedness" is important.

For each individual observer, I have performed three analyses of variance:

- an overall 3-way between-observers ANOVA with factors B/R (2 levels), S/U (2 levels) and nw (5 levels), that examines the entirety of the data for that observer; this ANOVA is the per-observer equivalent of the global ANOVA done in section 7.3.1.4, i.e. it contains data for all conditions all notch widths, but only for the current observer rather than data averaged across all observers;
- a 2-way between-observers ANOVA with factors S/U (2 levels) and nw (5 levels), that only looks at the data obtained under the B condition; and
- a 2-way mixed-design ANOVA with factors S/U (2 levels) and nw (5 levels), that only looks at the data obtained under the R condition.

The reason why I chose to do analyses separately for the B and R conditions has to do with the design of the experiment: a randomised (R) session would measure thresholds for all five different notch widths, while a blocked (B) session would only do so for one single value of notch width. This meant that, within an R-type session, there was a high degree of "relatedness" between the different notch widths, unlike within a B-type session, where different notch widths would be measured in different sessions. This meant that "notch width" would be a *within*-observer[31] factor in an ANOVA that is ran just for the R case, and a *between*-observers factor in an ANOVA ran just for the B case.

On the other hand, since blocks of five sessions ("sittings") were all ran under either the S or the U condition, this means that different S and U conditions were always obtained in different sessions. So the difference between S and U is the same – in terms of "relatedness" – as the difference between the notch widths, and consequently, the S vs U factor was a *between*-observers factor for both the "R" ANOVA and the "B" ANOVA.

In the ANOVAs for the B condition and for the R condition (the last two described in the bulleted list above), the nw factor was in one case *between*- and in the other case *within*-observers; but for the overall 3-way ANOVA (the first element in the list), this distinction cannot be made, and therefore everything is treated as a *between*-observers factor, ignoring the extra relatedness in some conditions. This can be seen as rather conservative, and it implies that, if a main effect of nw is found, then its associated significance level is reliable (and, if anything, underestimated); in other words, there is no danger of overestimating the significance of the effects of the nw factor.

Summing up,

[31] I have used the standard terminology for describing experiment design: *between-observers* and *within-observer* (or: *between-subjects* and *within-subject*). However in this case, an observer is really a *session*, so the term *between-sessions* would be more appropriate; I will, however, continue to use the standard terms *between-observers* and *within-observer*.

- the overall ANOVA has three between-observers factors (B/R, S/U and nw) and is therefore a between-observers ANOVA,
- the "B" ANOVA has two between-observers factors (S/U and nw) and is therefore a between-observers ANOVA, and
- the "R" ANOVA has one between-observers factor (S/U) and one within-observer factor (nw) and is therefore a mixed-design ANOVA.

The three different ANOVAs done for each of the three observers mean a total of 9 ANOVAs were produced, which I will discuss below.

For observer TP,

1) The overall ANOVA produced the following results:
   - Levene's test of equality of error variances was not significant, which means that the homogeneity of variance assumption was met;
   - There was a main effect of nw, $F(4, 80) = 52.94$, $p<.001$, which means that for this observer (as for all the others), threshold values averaged across all sets of conditions were significantly influenced by the notch width for which they were measured.
     The improvement in performance with increasing notch width is confirmed by a linear trend analysis ($F(4, 80) = 52.94$, $p<0.001$);
   - There were no significant main effects of S/U and B/R;
   - The interaction effect B/R*nw was not significant, although it was very close to significance: $F(4, 80) = 2.45$, $p = .053$;
   - There were no significant interaction effects of S/U*B/R, S/U*nw and S/U*B/R*nw.
2) The "B" ANOVA produced the following results:
   - Levene's test of equality of error variances was not significant, which means that the homogeneity of variance assumption was met;
   - There was a main effect of nw, $F(4, 40) = 27.36$, $p<.001$, which means that for this observer (as for all the others), threshold values averaged across the S and U conditions were significantly influenced by the notch width for which they were measured;
   - There was no main effect of S/U;
   - There was no interaction effect of S/U*nw.
3) The "R" ANOVA produced the following results:
   - Mauchly's test of sphericity was not significant, which means that the sphericity assumption relating to the within-observer factor (nw) has been met and therefore there is no need for a correction;
   - Levene's test of *equality of error variances* was not significant for any of the levels of the within-observer factor (nw) with the exception of the first one, which means that I should be cautious in interpreting the results, since one of the assumptions of the analysis is violated;
   - There was a main effect of nw, $F(4, 32) = 81.80$, $p<.001$, which means that for this observer (as for all the others), threshold values averaged across the S and U conditions were significantly influenced by the notch width for which they were measured;

- There was no main effect of S/U;
- There was no interaction effect of nw*S/U.

For observer HS,

1) The overall ANOVA produced the following results:
   - Levene's test of equality of error variances was not significant, which means that the homogeneity of variance assumption was met;
   - There was a main effect of nw, $F(4, 80) = 102.42$, $p<.001$, which means that for this observer (as for all the others), threshold values averaged across all sets of conditions were significantly influenced by the notch width for which they were measured.
     The improvement in performance with increasing notch width is confirmed by a linear trend analysis ($F(4, 80) = 102.42$, $p<0.001$);
   - There were no significant main effects of S/U and B/R, although the main effect of B/R was close to significance: $F(4, 80) = 3.12$, $p = .081$; this means that the S vs U condition did not significantly influence threshold measurements, while the B vs R condition did so marginally;
   - There were no significant interaction effects of S/U*B/R, S/U*nw, B/R*nw and S/U*B/R*nw.
2) The "B" ANOVA produced the following results:
   - Levene's test of equality of error variances was not significant, which means that the homogeneity of variance assumption was met;
   - There was a main effect of nw, $F(4, 40) = 58.40$, $p<.001$, which means that for this observer (as for all the others), threshold values averaged across the S and U conditions were significantly influenced by the notch width for which they were measured;
   - There was no main effect of S/U;
   - There was no interaction effect of S/U*nw.
3) The "R" ANOVA produced the following results:
   - Mauchly's test of sphericity was not significant, which means that the sphericity assumption relating to the within-observer factor (nw) has been met and therefore there is no need for a correction;
   - Levene's test of equality of error variances was not significant for any of the levels of the within-observer factor (nw), which means that the homogeneity of variance assumption was met;
   - There was a main effect of nw, $F(4, 32) = 100.38$, $p<.001$, which means that for this observer (as for all the others), threshold values averaged across the S and U conditions were significantly influenced by the notch width for which they were measured;
   - There was no main effect of S/U;
   - There was no interaction effect of nw*S/U.

For observer WS,

1) The overall ANOVA produced the following results:

- Levene's test of equality of error variances was not significant, which means that the homogeneity of variance assumption was met;
- There was a main effect of nw, $F(4, 80) = 20.28$, $p<.001$, which means that for this observer (as for all the others), threshold values averaged across all sets of conditions were significantly influenced by the notch width for which they were measured.
  The improvement in performance with increasing notch width is confirmed by a linear trend analysis ($F(4, 80) = 20.28$, $p<0.001$);
- There was a main effect of S/U, $F(1, 80) = 4.12$, $p<.05$;
- There were no significant main effects of S/U and B/R;
- There were no significant interaction effects of S/U*B/R, S/U*nw, B/R*nw and S/U*B/R*nw.

2) The "B" ANOVA produced the following results:
   - Levene's test of equality of error variances was not significant, which means that the homogeneity of variance assumption was met;
   - There was a main effect of nw, $F(4, 40) = 14.03$, $p<.001$, which means that for this observer (as for all the others), threshold values averaged across the S and U conditions were significantly influenced by the notch width for which they were measured;
   - There was no main effect of S/U;
   - There was no interaction effect of S/U*nw.
3) The "R" ANOVA produced the following results:
   - Mauchly's test of sphericity was not significant, which means that the sphericity assumption relating to the within-observer factor (nw) has been met and therefore there is no need for a correction;
   - Levene's test of equality of error variances was not significant for any of the levels of the within-observer factor (nw), which means that the homogeneity of variance assumption was met;
   - There was a main effect of nw, $F(4, 32) = 27.46$, $p<.001$, which means that for this observer (as for all the others), threshold values averaged across the S and U conditions were significantly influenced by the notch width for which they were measured;
   - There was no main effect of S/U;
   - There was no interaction effect of nw*S/U.

7.3.2 Learning analysis

7.3.2.1 Graphical representation of data

It helps at this point to recapitulate the design of the experiment – data (i.e. contrast thresholds) were obtained as a function of notch width. Under each set of conditions, my observers obtained repeated estimates (five estimates, to be more precise) at every notch width.

An interesting visual exploration of the data is to plot the thresholds for the different notch widths as a function of estimate number. I do this in Figure 29 below, which plots the evolution of threshold estimates under one of the four sets of condition (UB), and does this for each of the three observers. Error bars are shown for each data

point, and indicate ±1 standard error (SE) across sessions. In order to make the plots easier to grasp, I have only included the two extreme values of notch width (0.0 octaves and 2.0 octaves), which are enough to make the case that learning occurs for large notch widths.

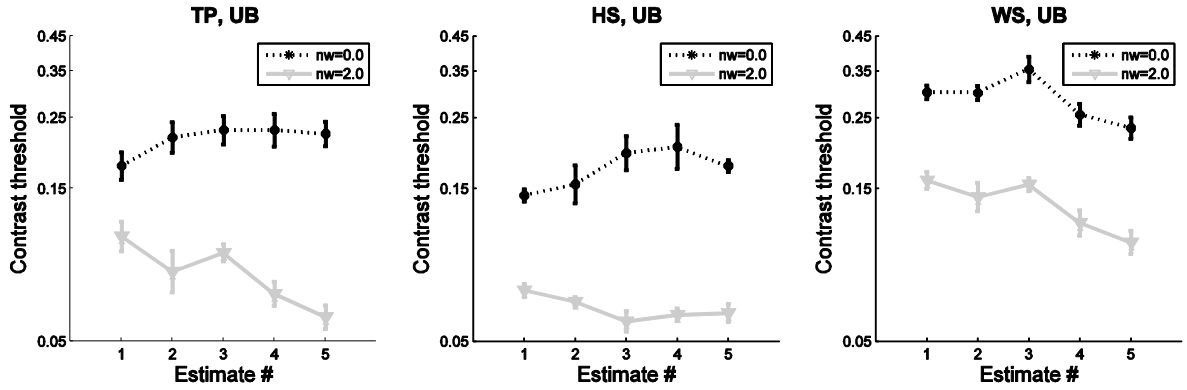

Figure 29: Thresholds obtained across the different estimates, for the two extreme values of notch width, 0.0 octaves (black) and 2.0 octaves (grey). The different plots are for each of the three observers, under noise condition "unique" and presentation condition "blocked". Error bars are shown for each data point, and indicate ±1 standard error (SE) across sessions.

Looking at Figure 29 above, it seems that the threshold estimates obtained late in the practice, when observers had gained experience performing the task, are generally lower than those obtained early in the practice. This is particularly true for the larger notch width (2.0 octaves).

The same improvement in performance can be looked at from a different angle, as in Figure 30, which shows the variation of threshold as a function of notch width for each of the five different estimates. Figure 30 represents the same data as Figure 29, with the difference that in the former, notch width describes the abscissa and the estimate number is the parameter differentiating curves, while in the latter, the two roles are reversed (i.e. the estimate number is on the abscissa and the notch width is the parameter).

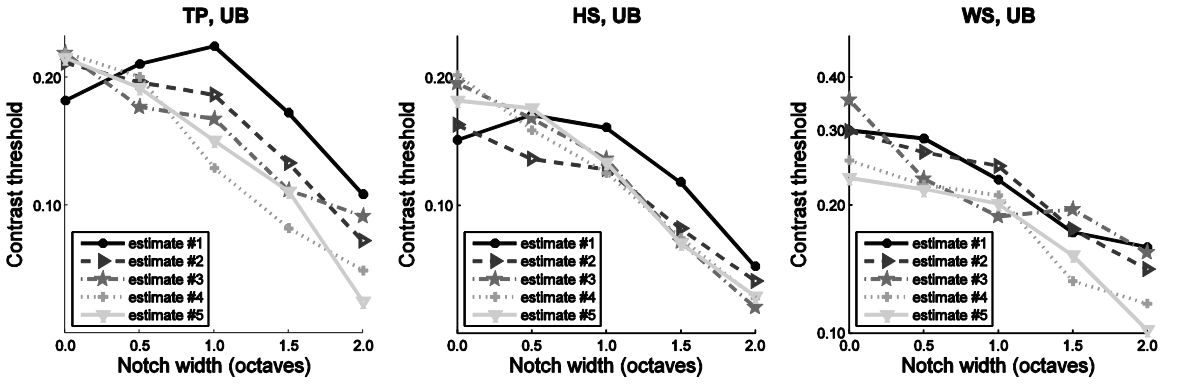

Figure 30: Plots comparing the different results obtained by each observer for each of the five estimates, under condition set UB. Lighter shades of grey represent late estimates, which, as can be seen, generally produce lower thresholds.

It can be seen in Figure 30 above that the curves corresponding to early estimate are generally lower than the ones corresponding to late estimate. There is also, of course,

a general decreasing trend for all curves (i.e. lower thresholds for wider notches, as expected).

The previous two graphs (Figure 29 and Figure 30) create the expectation that there will be a main effect of estimate number. To test this formally, I did the analysis of variance described in the following section.

7.3.2.2ANOVA per set of conditions

It is clear that improvement in performance (as measured by decreasing threshold) over repeated estimates is an indication of the rate of learning. This can be seen in the vertical shift that occurs between the different estimates' curves in Figure 30. However, in the same figure, a change in *shape* between those curves can also be seen, which seems to suggest that the rate of decrease depends on notch width (i.e. learning is different for different notch widths). This apparent trend would be confirmed by a significant interaction effect between notch width and estimate number, as part of an ANOVA[32]. It is with this purpose in mind that I performed four 2-way repeated-measures ANOVAs, with factors notch width (5 levels), and estimate number (5 levels)[33]. Each of those four ANOVAs correspond to the four possible combinations of my two conditions: SB, SR, UB and UR, and they ask the question of whether estimate number influences threshold, whether notch width influences threshold, and whether an interaction between the two affects thresholds.

It is worth mentioning that, since measurements were taken from only three observers, Mauchly's test of sphericity could not be calculated for any of the four ANOVAs. I will, however, assume sphericity in their interpretation which is given below.

The "SB" ANOVA:

- There was a significant main effect of notch width, $F(4, 8) = 28.88$, $p<.001$, which suggests that, in general, threshold values are influenced by notch width; looking at the means, there is a tendency for threshold to *decrease* as nw *increases*, as one would expect from the current literature.
- There was a significant main effect of estimate number, $F(4, 8) = 36.43$, $p<.001$, which shows that learning does indeed exist (i.e. late thresholds are lower than early thresholds), as can be seen in the graphs.
- There was a significant interaction effect between the two factors (nw*en), $F(16, 32) = 2.35$, $p<.05$, which indicates that the different estimates had different effects on the log threshold values depending on the notch width of the noise.

The "SR" ANOVA:

[32] Within-observer (repeated-measures) ANOVAs are usually done with more than three observers, however in this case this analysis provided important information about the rate of learning. The only downside to having only three observers for this ANOVA was that Mauchly's test of sphericity could not be calculated, as is reported further below.

[33] As I have done previously, I refer to these factors in shorthand – nw and en.

- There was a significant main effect of notch width, F(4, 8) = 11.89, p<.005, which suggests that, in general, threshold values are influenced by notch width; looking at the means, there is a tendency for threshold to *decrease* as nw *increases*, as one would expect from the current literature.
- There was a significant main effect of estimate number, F(4, 8) = 11.00, p<.005, which shows that learning does indeed exist (i.e. late thresholds are lower than early thresholds), as can be seen in the graphs.
- The interaction effect between the two factors (nw*en) was not significant, F(16, 32) = 1.95, p>.05, which indicates that the different estimates had roughly the same effect on the log threshold values regardless of notch width.

The "UB" ANOVA:

- There was a significant main effect of notch width, F(4, 8) = 102.44, p<.001, which suggests that, in general, threshold values are influenced by notch width; looking at the means, there is a tendency for threshold to *decrease* as nw *increases*, as one would expect from the current literature.
- There was a significant main effect of estimate number, F(4, 8) = 4.64, p<.05, which shows that learning does indeed exist (i.e. late thresholds are lower than early thresholds), as can be seen in the graphs.
- There was a significant interaction effect between the two factors (nw*en), F(16, 32) = 2.46, p<.05, which indicates that the different estimates had different effects on the log threshold values depending on the notch width of the noise.

The "UR" ANOVA:

- There was a significant main effect of notch width, F(4, 8) = 11.78, p<.005, which suggests that, in general, threshold values are influenced by notch width; looking at the means, there is a tendency for threshold to *decrease* as nw *increases*, as one would expect from the current literature.
- There was a significant main effect of estimate number, F(4, 8) = 14.01, p<.005, which shows that learning does indeed exist (i.e. late thresholds are lower than early thresholds), as can be seen in the graphs.
- The interaction effect between the two factors (nw*en) was not significant, F(16, 32) = 1.54, p>.05, which indicates that the different estimates had about the same effect on the log threshold values regardless of notch width.

7.3.2.3 Regression analyses

It would seem, from the *threshold of contrast* as a function of *estimate number* plots (Figure 29), that the rate of learning increases with increasing notch width (i.e. the grey line has a more accentuated downward slope than the black line). However to formalise this in statistical terms, I have decided to do a regression analysis.

The first step was to do a regression analysis for each combination of observer, S/U condition, B/R condition and notch width. This meant a total of 3 x 2 x 2 x 5 = 60

regressions, which produced 60 pairs of regression line parameters (slope and y-intercept). The slopes represent units of contrast that the threshold goes down by from one estimate to the next.

What needed to be done at this point, in order to answer the research question posed at the beginning of the vision experiment chapter was to plot, for each combination of observer, S/U condition and B/R condition, the rates of learning (represented by the regression line slopes) as a function of the notch width for which they had been calculated. This meant a total of 12 graphs, for the 3 x 2 x 2 possible combinations of observer, S/U condition and B/R condition.

Finally, a regression was done for each of those plots; thus, the obtained b_1 coefficients effectively represented the change in the rate of learning with each 0.5-octave increase of the notch width, for each observer and for each set of conditions. The results are presented in Table 5 below:

Table 5: Results of the regression analyses done to find the different slopes of learning manifested by the three observers under the four possible sets of conditions. The coefficient of determination R^2 (where R is the Pearson product-moment correlation coefficient) is also given for each case.

Observer	S/U	B/R	b_1	significance of b_1	R^2
TP	S	B	0.020	0.016	0.8918
TP	S	R	0.008	0.270	0.3763
TP	U	B	0.011	0.128	0.5926
TP	U	R	0.004	0.508	0.1577
HS	S	B	0.007	0.060	0.7437
HS	S	R	0.012	0.026	0.8503
HS	U	B	0.000	0.934	0.0027
HS	U	R	0.005	0.103	0.6412
WS	S	B	0.018	0.006	0.9432
WS	S	R	0.003	0.296	0.3481
WS	U	B	0.003	0.524	0.1465
WS	U	R	0.011	0.126	0.5976

As can be seen from Table 5 above, as well as from the plots in the next three figures, most of the b_1 coefficients were greater than zero, indicating that the rate of learning was higher for larger notch widths.

As regards the significance values of these regression coefficients, it can be seen that, for two of my three observers (TP, WS), the only set of conditions for which the slope was significant was the SB condition; under this condition, the third observer (HS) was not far from significance, with a p-value of .060, but only reached significance in the SR condition.

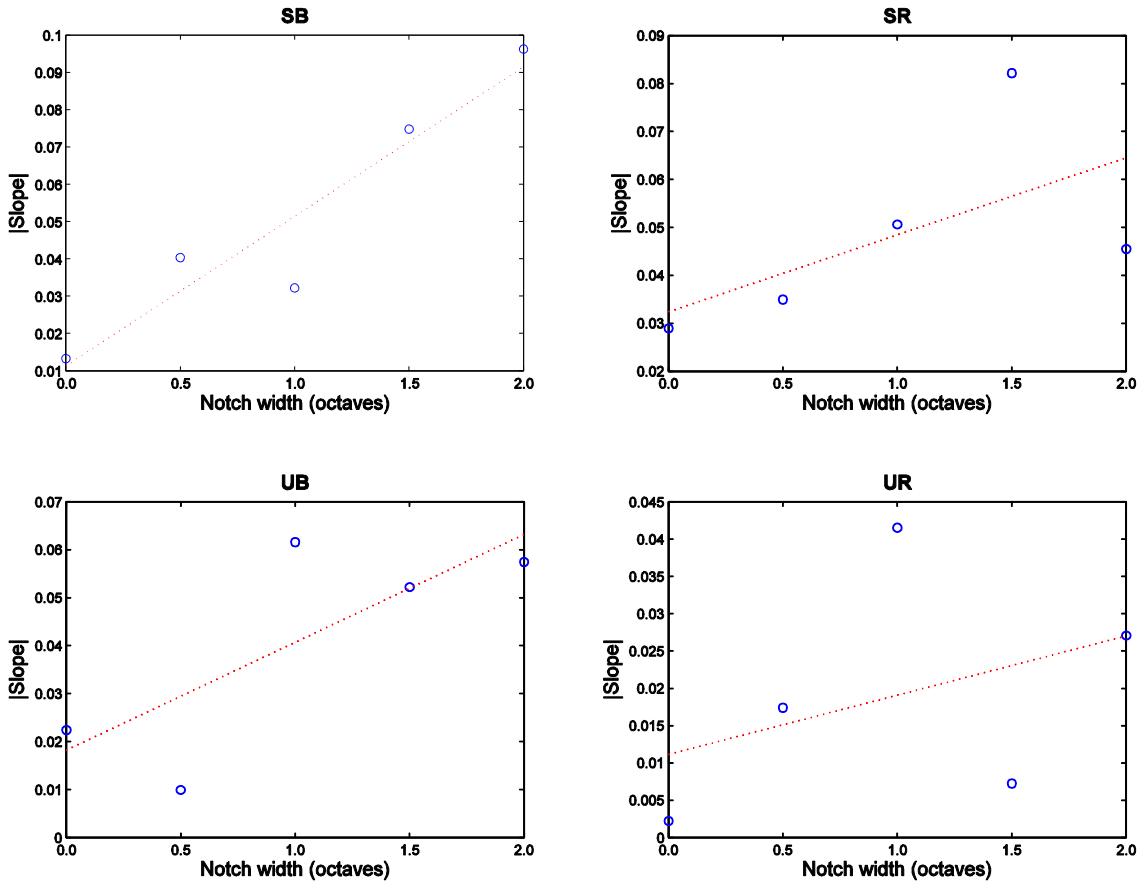

Figure 31: The plots corresponding to each of the four regressions done for observer TP to find out the "slope of learning" (represented by value b_1 from Table 5 above).

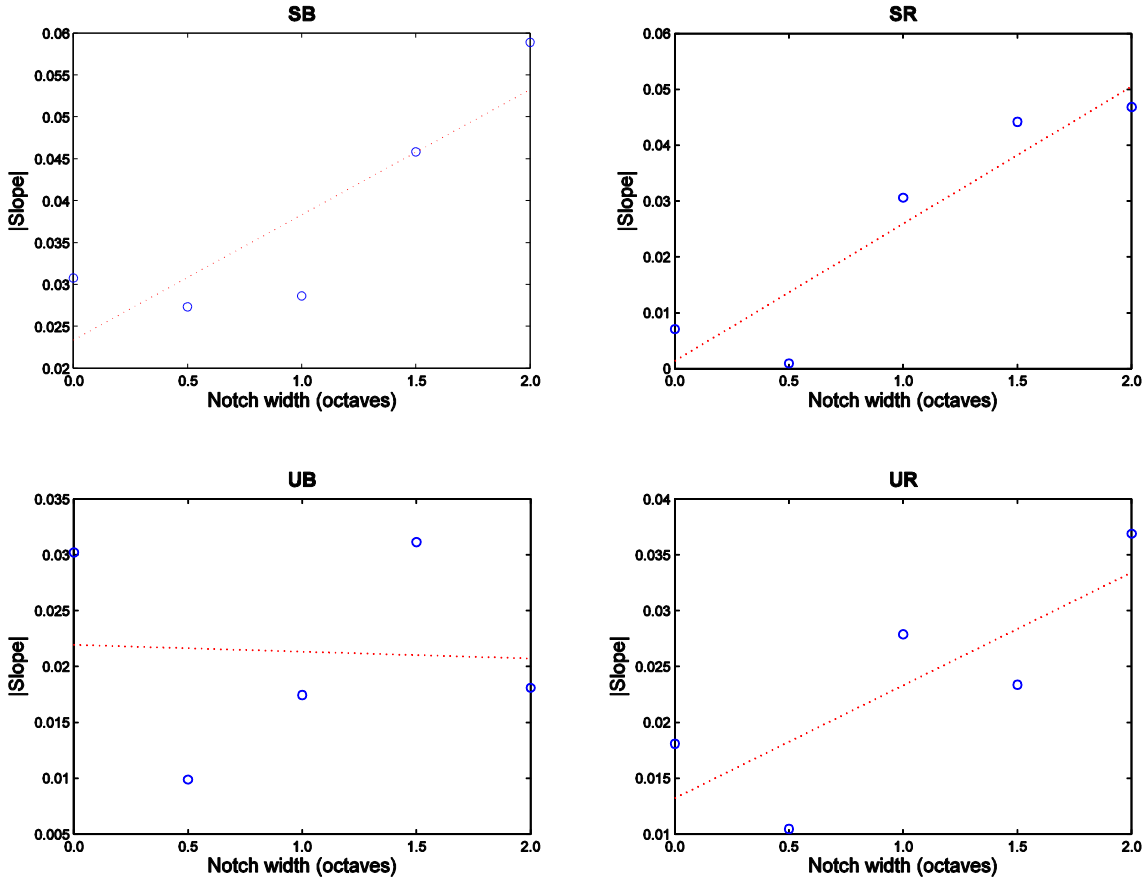

Figure 32: The plots corresponding to each of the four regressions done for observer HS to find out the "slope of learning" (represented by value b_1 from Table 5 above).

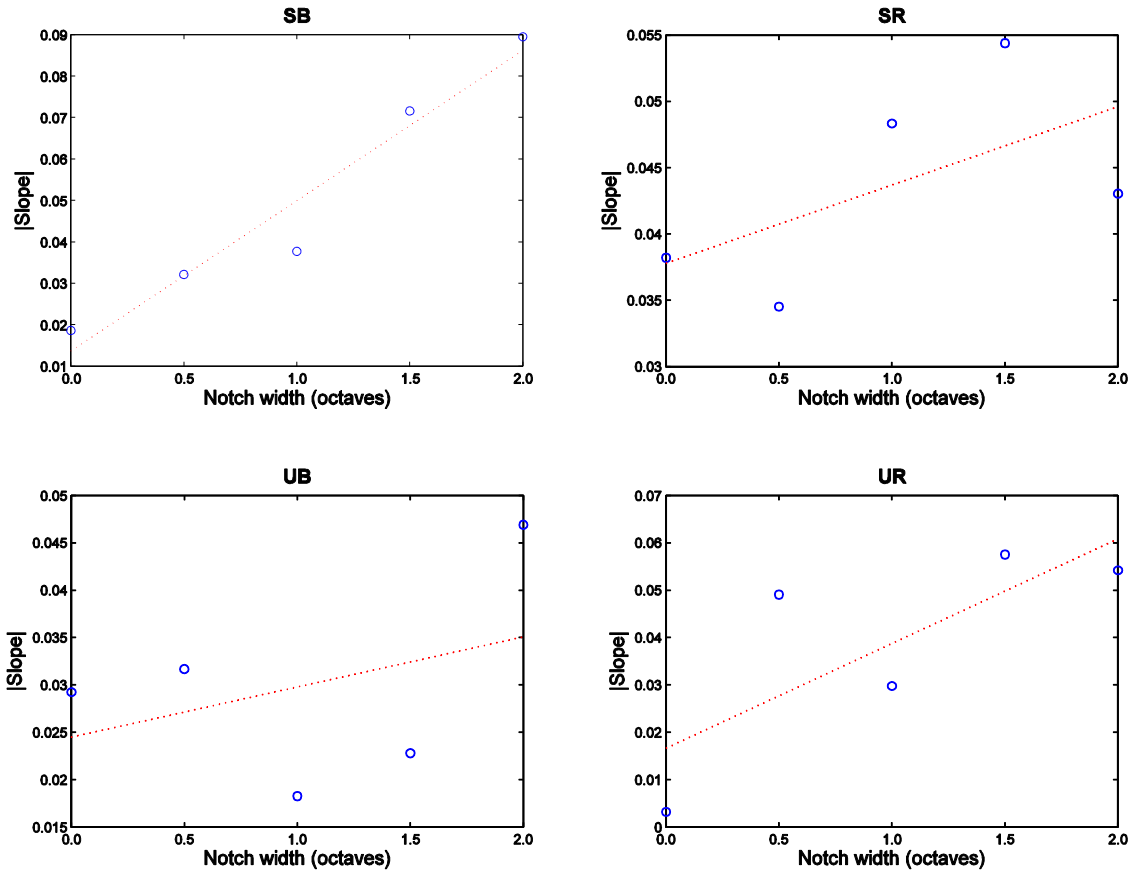

Figure 33: The plots corresponding to each of the four regressions done for observer WS to find out the "slope of learning" (represented by value b_1 from Table 5 above).

7.4. Discussion

7.4.1 Main analysis

7.4.1.1 Effect of "blocked" vs "randomised"

As I have reported in the Results (section 7.3.1.4), the overall ANOVA showed there was a main effect of B vs R, with B producing better overall performance. If one were to run notched-noise experiments, this main effect would translate to different channel *sensitivities* being inferred when using blocked design in comparison to when using randomised design (channel *shapes* would not be different, though, just because of this main effect). Therefore, if one wishes to make such a measurement in which sensitivity is relevant, attention must be paid to the type of design (blocked or randomised) that is used.

On the other hand there was no significant interaction effect between B vs R and notch width, which means that, first of all, the *shape* of the threshold as a function of notch width function is fixed and, consequently, that the shape of the inferred filter is fixed and would therefore not be affected by the choice of either a blocked or a randomised design. This means that inferences about channel bandwidth done in the classical literature – for example, in (Stromeyer III & Julesz, 1972) – are "safe", in the sense that their results are not likely to have been influenced by their – perhaps inadvertent – choice of *randomised* design or, on the contrary, of the "default" *blocked* design.

The overall improvement in the B case can be thought of as a change in gain, and, as shown in the Dosher and Lu papers (1998, 1999), the cause of this change is likely to be either (target) signal enhancement or noise reduction. Since in the present case, the target signal was the same for both B and R conditions, and my observers had the same amount of practice with the targets in these two conditions, it follows that the more likely mechanism is that of noise reduction. The noise reduction mechanism (i.e. learning the properties of the noise and how to discount it) cannot occur on the basis of a single trial, but rather across the time scale over which B and R differ – which is that of a *group of sessions*.

7.4.1.2 Effect of "same" vs "unique"

As also reported in section 7.3.1.4, the S vs U condition did not show a significant main effect. The most important implication of such a result is that it goes against the predictions of the ideal observer model, as was also the case, for example, in Watson et al.'s paper (1997), in which a significant difference between those two conditions was also not found.

In the previous section, I have found an advantage for blocked conditions, indicating that noise reduction can be achieved over the time course of a sitting. However, the fact that I found no difference between "same" and "unique" indicates that noise cannot be discounted over the time course of a single trial. I remind the reader that observers did not know when the noises in the two intervals were the same and when they were different.

Although S vs U did not show a significant main effect, there was a significant interaction effect between this factor and notch width (S/U*nw). This means that, in contrast to the B/R condition, the S/U condition *can* invalidate inferences about channel shape, if this factor has not been controlled. In other words, from an "S/U" perspective, such inferences are *not* "safe".

A more subtle point to make here is based on the idea that two factors (I am referring to S/U and nw) may show a significant interaction effect between them (S/U*nw), but they may also have equal means, such that neither overall main effect would turn out to be significant. However, this does not imply that all pairs of levels from the two factors are equal. So even though there is no overall difference, the difference might exist only for particular levels. This would potentially be revealed by an analysis of simple effects, which might reveal that, only for a particular notch width, there *is* a difference between the two levels of the S/U factor. This in itself might explain some of the conflicting results in the literature that have inconsistently found differences between "same" and "unique" (Ahumada Jr & Beard, 1997; Eckstein, et al., 1997a; Swift & Smith, 1983; Watson, et al., 1997).

7.4.2 Learning effects

The four ANOVAs done per set of conditions (section 7.3.2.2 for details) revealed that the only times there was a significant interaction between notch width (nw) and estimate number (en) was in condition sets SB and UB – in other words, whenever the

presentation condition was blocked. So it might be said that learning properties of noise with a particular notch width is easier when that notch width is "blocked".

In all four ANOVAs, threshold depended on notch width – it would have been surprising if it had not, given the existing literature – but it also depended on estimate number. Although ANOVA is not able to tell the "direction" of this dependence, by simply looking at the means it can be seen that thresholds go *down* with practice (i.e. with estimate number). Therefore, in all cases, there is significant learning, but even more interestingly, the influence of practice is different for different notch widths, at least in the B conditions.

Once these ANOVAs established that learning does occur, the regression analysis done in the following section (7.3.2.3) allowed for a closer look at the progression of learning. Specifically, it showed significant positive slopes (i.e. increases in the rates of learning with increasing notch widths) for two observers under the SB condition, and a close-to-significance slope for the third observer, under the same SB condition.

This idea is related to one of Watson et al.'s (1997) conclusions, which was that "simplicity" of the noise enables learning. The highest potential for such learning due to simplicity occurs during "blocked" conditions; of course, even in the "randomised" conditions, larger notch widths are simpler (just because they leave fewer frequency components in the noise) – but they are more difficult to learn because of the randomisation. So large notches might have an effect of reducing the SNR in the target channel and reducing the degrees of freedom in the noise, but it may also make the noise easier to learn with extended practice in blocked conditions.

8. HEARING EXPERIMENT

8.1. Introduction

The channel model in vision can be considered as having been inspired by the auditory-frequency-selective channels in hearing. Despite evidence that, in hearing, these channels have an early physiological basis and therefore are likely to be hard-wired, one still cannot completely rule out the contribution of later stages to those auditory channels; please refer back to section 3.6 where I discussed the debate in the hearing literature on physiological (low-level) filter measurements vs psychophysical (high-level) ones.

It is, therefore, of interest to ask the question "to what degree are the channels hard-wired?" in hearing as well, as I already have done in vision, and I will pursue this question with the help of a similar experimental paradigm.

8.2. Methods

8.2.1 Observers

There was only one observer used for the hearing experiment – Tudor Popescu (TP). Data collection was started, but not finished at the time of submission of the thesis, for a second observer – Lucia Ferrara (LF), whose data are not taken into consideration here.

8.2.2 Task

The structure of a trial was the same as in the vision experiment, i.e. two intervals containing *target+noise* and *noise* alone (in a random order), separated by a blank. The duration of the blank was the same as in vision – 500 ms, however the duration of the intervals was 250 ms, compared to 50 ms in vision. In the vision case, the blank was a uniform field of 50% grey luminance displayed for 500 ms. In hearing, the blank simply consisted of 500 ms of silence.

8.2.3 Equipment

Both the target and the noise were played out through the right channel of a mobile 24-bit M-Audio Transit external (USB-connected) sound card. The signal was then sent – via a balanced line – to a final amplifier, from where it was presented monaurally to the right ear of the observer via Etymotic ER2 insert earphones. The earphones had been calibrated to have a flat response at the eardrum.

The sound card was controlled by means of a Core Audio library, which was originally written in C++ but was compiled as a MEX file in order to be called through MATLAB. The computer used was a Mac Mini.

8.2.4 Calibration

The Etymotic ER2 insert earphones had been factory calibrated to have a flat response at the eardrum.

8.2.5 Procedure

The trials were displayed using the same staircase procedure as the one used in the vision experiment (described in detail in section 7.2.5), also used with six reversals. Just as in vision, I used all staircases in conjunction with the same performance level of 75%.

In vision, I had found that starting each staircase with an initial contrast of 40% leads so the shortest convergence time for the staircase. Coincidentally (or not), I have found that the same percentage (which in hearing refers to *sound level*) also minimises staircase duration in the hearing experiment. Therefore, a 40% level was used to start each staircase.

8.2.6 Conditions

The conditions were the same as in the vision experiment, namely notch width, noise condition (S/U) and presentation condition (B/R). The order that the sessions were ran in was also the same as in vision (see Appendix A for a full outline, as well as Appendix B for a note on the counter-balancing of trials).

The only difference was in the values of the notch widths that I used. Instead of using octave values of 0.0, 0.5, 1.0, 1.5 and 2.0, as I had in vision, here I used notch widths of 0 Hz, 200 Hz, 400 Hz, 600 Hz and 800 Hz. These notch widths, expressed as the difference from the target frequency of 2,000 Hz, are equivalent to 0.0000, 0.1444, 0.2895, 0.4361 and 0.5850 octaves. The reason why I chose to use smaller step sizes than in vision, and spaced *linearly* rather than *logarithmically*, was that in hearing, the channels are more "squashed in" than in vision, which would have lead to a complete obliteration of the channel had I used large notch widths such as those above one octave.

8.3. Results

Throughout this Results section, I will plot thresholds using the *logarithm* of the sound level values rather than the sound level values themselves. I use the logarithm values rather than the plain values for the same reason why I also did this in the vision

case (see section 7.3). I will therefore be plotting "log level" values of threshold, but, as in vision, will label the y-axes in terms of plain values for easier interpretation.

I will also continue to make use of the shorthand notation I have defined for the experimental conditions ("S/U" for "same/unique", "B/R" for "blocked/randomised") and for variables that will appear as factors in analyses of variance ("nw" for notch width and "en" for estimate number).

In the hearing experiment, I measured the sound level required to detect a target stimulus embedded in noise, as a function of the notch width in that noise. As I did in the vision case, I will start by visually representing (plotting) the data, and then do the statistics that are able to test my main hypothesis, as well as emphasise any perceptual learning effects.

8.3.1 Main analysis

My initial hypotheses for the hearing experiment were similar to the ones for the vision experiment:

- whether there would be a main effect of the B as opposed to the R condition, and whether B/R would interact with nw;
- whether there would be a main effect of the S as opposed to the U condition, and whether S/U would interact with nw;
- whether the B/R factor would be most apparent in the U condition, because presumably under the S condition, there is no "room" for further learning.

Please see the last paragraph of section 7.3.1.4 for an explanation regarding the normality assumption required by parametric statistical analyses, which applies to the vision experiment as well as to the hearing experiment.

8.3.1.1 Graphical comparison between "blocked" and "randomised"

The most straightforward way to compare the B and R conditions is to plot on the same graph the curves representing threshold as a function of notch width obtained for each of them. Figure 34 below does this for observer TP (the only one for this experiment), separately for the S and U conditions. Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.

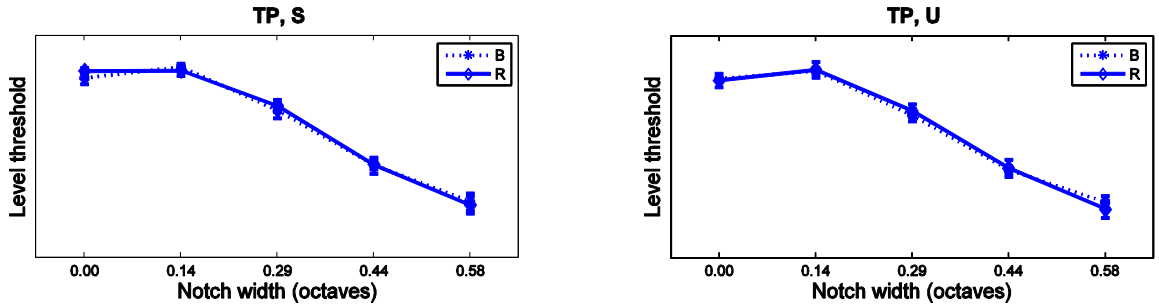

Figure 34: Graphs comparing results for the B and R conditions. The six graphs are for each of the three observers and under each S/U condition. Each data point represents the average of the five different estimates that were taken for that observer, under that set of conditions, and for that value of notch width. Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.

As can be seen from the plots above, the results for the B and R conditions seem to be almost identical, both for the S and for the U conditions. This will be confirmed later, in section 8.3.1.3.

8.3.1.2 Graphical comparison between "same" and "unique"

An overall picture of the difference between the S and U noise conditions can be gathered by plotting thresholds (averaged across all estimates) as a function of notch width, for two different curves that represent the S and U conditions. There are two such graphs, for the two presentation conditions (B and R), and they are shown in Figure 35 below. Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.

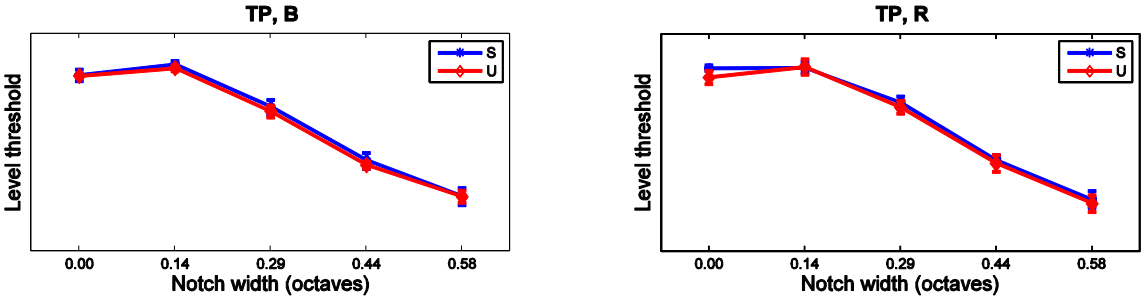

Figure 35: Graphs comparing results for the S and U conditions. The two plots correspond to the B and R conditions. Each data point represents the average of the five different estimates that were taken under that set of conditions and for that value of notch width. Error bars are shown for each data point, and indicate ±1 standard error (SE) across estimates.

Here too it seems that there is almost no difference at all between the results obtained in the S condition and those obtained in the U condition, in both the B and R cases. Again, though, I will let the statistics have the final word, in section 8.3.1.3.

8.3.1.3 Analyses per observer

After visually exploring the data using plots, I have performed a series of ANOVAs analogous to the ones in section 7.3.1.5, to test whether the independent variables that I manipulated in this experiment influenced thresholds.

For observer TP (the only one who collected data for the hearing experiment), I have performed three analyses of variance:

- an overall 3-way between-observers ANOVA with factors B/R (2 levels), S/U (2 levels) and nw (5 levels), that examines the entirety of the data for that observer;
- a 2-way between-observers ANOVA with factors S/U (2 levels) and nw (5 levels), that only looks at the data obtained under the B condition; and
- a 2-way mixed-design ANOVA with factors S/U (2 levels) and nw (5 levels), that only looks at the data obtained under the R condition.

The results of the three ANOVAs are presented below. Please see the corresponding results section for the vision experiment (7.3.1.5) for a discussion about the nature of these three ANOVAs.

1) The overall ANOVA produced the following results:
   - Levene's test of equality of error variances was not significant, which means that the homogeneity of variance assumption was met;
   - There was a main effect of nw, $F(4, 80) = 345.37$, $p<.001$, which means that for this observer, threshold values averaged across all sets of conditions were significantly influenced by the notch width for which they were measured;
     The improvement in performance with increasing notch width is confirmed by a linear trend analysis ($F(4, 80) = 345.37$ $p<0.001$);
   - There were no significant main effects of S/U and B/R
   - There were no significant interaction effects of B/R*nw, S/U*B/R, S/U*nw and S/U*B/R*nw.
2) The "B" ANOVA produced the following results:
   - Levene's test of equality of error variances was not significant, which means that the homogeneity of variance assumption was met;
   - There was a main effect of nw, $F(4, 40) = 197.09$, $p<.001$, which means that for this observer, threshold values obtained in the B condition and averaged across the S/U condition were significantly influenced by the notch width for which they were measured;
   - There was no main effect of S/U;
   - There was no interaction effect of S/U*nw.
3) The "R" ANOVA produced the following results:
   - Mauchly's test of sphericity was not significant, which means that the sphericity assumption relating to the within-observer factor (nw) has been met and therefore there is no need for a correction;
   - Levene's test of equality of error variances was not significant for any of the levels of the within-observer factor (nw);
   - There was a main effect of nw, $F(4, 32) = 502.66$, $p<.001$, which means that for this observer, threshold values obtained in the R condition and averaged across the S/U condition were significantly influenced by the notch width for which they were measured;
   - There was no main effect of S/U;
   - There was no interaction effect of nw*S/U.

8.3.2 Learning effects

8.3.2.1 Graphical representation of data

I will shortly remind the reader that the hearing experiment had a very similar design to the vision experiment; data (i.e. level thresholds) were obtained as a function of notch width. Under each set of conditions, observer TP obtained repeated estimates (five estimates, to be more precise) at every notch width.

It is interesting here as well to explore the data visually by plotting the thresholds for the different notch widths as a function of estimate number. In Figure 36 below I plotted the evolution of threshold estimates under one of the four sets of condition (UB). Error bars are shown for each data point, and indicate ±1 standard error (SE) across sessions. In order to make the plots easier to grasp, I have only included the two extreme values of notch width (0.00 octaves and 0.58 octaves), which are enough to make the case that learning occurs for large notch widths.

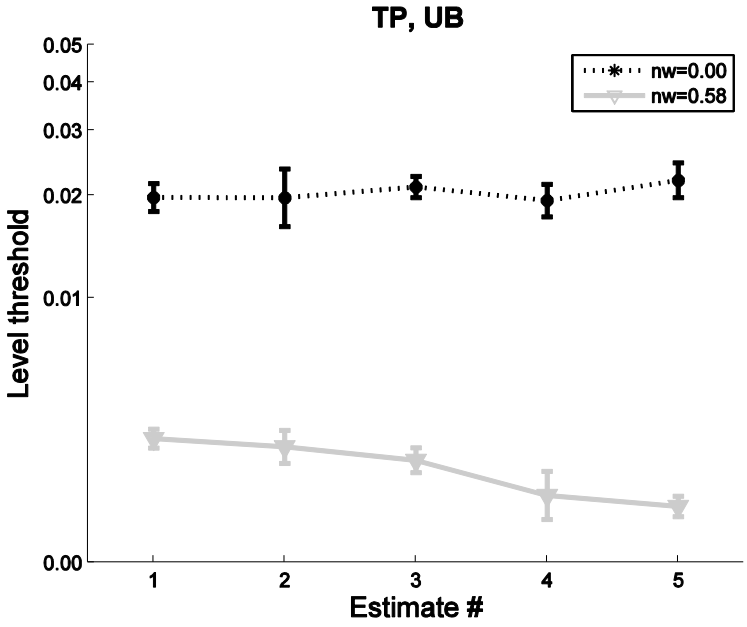

Figure 36: Thresholds obtained across the different estimates, for the two extreme values of notch width, 0.00 octaves (black) and 0.58 octaves (grey), under conditions UB. Error bars are shown for each data point, and indicate ±1 standard error (SE) across sessions.

Looking at Figure 36 above, it seems that the threshold estimates obtained late in the practice, when observers had gained experience performing the task, are generally lower than those obtained early in the practice. This is particularly true for the larger notch width (0.58 octaves).

The same improvement in performance can be looked at from a different angle, as in Figure 37 below, which shows the variation of threshold as a function of notch width for each of the five different estimates. Figure 37 represents the same data as Figure 36, with the difference that in the former, notch width describes the abscissa and the estimate number is the parameter differentiating curves, while in the latter, the two roles are reversed (i.e. the estimate number is on the abscissa and the notch width is the parameter).

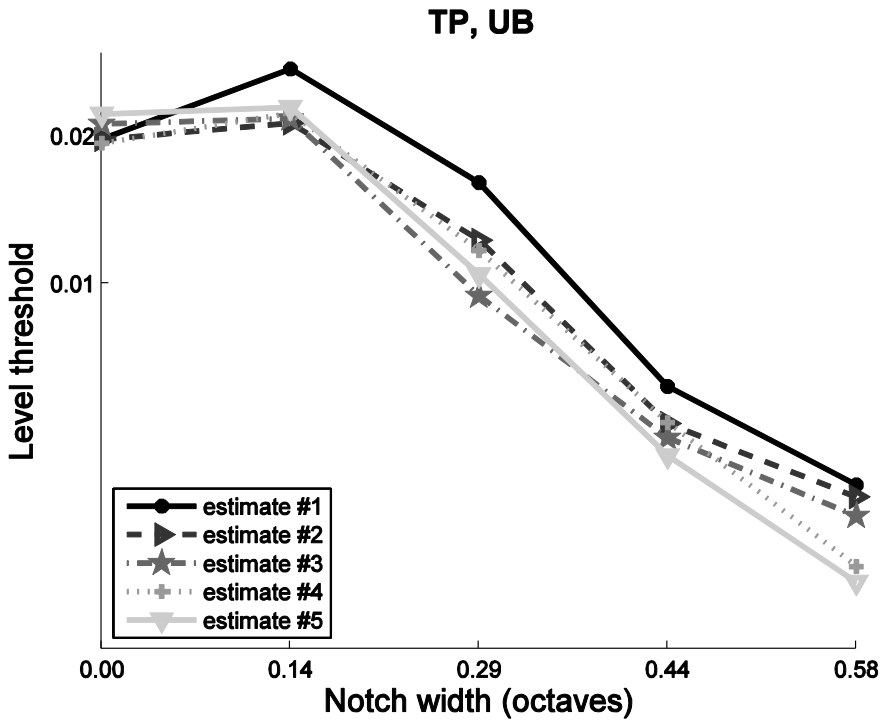

Figure 37: Plots comparing the different results obtained for each of the five estimates, under condition set UB. Lighter shades of grey represent late estimates, which, as can be seen, generally produce lower thresholds.

It can be seen in Figure 37 above that the curves corresponding to early estimate are generally lower than the ones corresponding to late estimate. There is also, of course, a general decreasing trend for all curves (i.e. lower thresholds for wider notches, as expected).

8.3.2.2Regression analyses

It would seem, from the *threshold of level* as a function of *estimate number* plots (Figure 36), that the rate of learning increases with increasing notch width (i.e. the grey line has more of a downward slope than the black line), however to formalise this in statistical terms, I have decided to perform a regression analysis, in the same way I have performed it in the corresponding section of the vision experiment (see section 7.3.2.3).

The first step was to do a regression analysis for each combination of observer, S/U condition, B/R condition and notch width. This meant a total of 1 x 2 x 2 x 5 = 20 regressions, which produced 20 pairs of regression line parameters (slope and y-intercept).

What needed to be done at this point, in order to answer the research question posed at the beginning of the hearing experiment chapter was to plot, for each combination of observer, S/U condition and B/R condition, the rates of learning (represented by the regression line slopes) as a function of the notch width for which they had been calculated. This meant a total of 4 graphs, for the 1 x 2 x 2 possible combinations of observer, S/U condition and B/R condition.

Finally, a regression was done for each of those 4 plots; thus, the obtained b_1 coefficients effectively represented the slope of learning, for each set of conditions. The results are presented in Table 6 below:

Table 6: Results of the regression analyses done to find the different slopes of learning manifested by observer TP under the four possible sets of conditions.

Observer	S/U	B/R	b_1	significance of b_1
TP	S	B	0.010	0.101
TP	S	R	0.012	0.067
TP	U	B	0.009	0.043
TP	U	R	0.009	0.177

As can be seen from Table 6 above, as well as from Figure 38 below, all of the b_1 coefficients (i.e. slopes of the regression lines) were greater than zero, indicating that the rate of learning was always higher for larger notch widths.

As regards the significance values of these regression coefficients, it can be seen that the slope was not very far from significance under all conditions, but only reached significance for the UB condition.

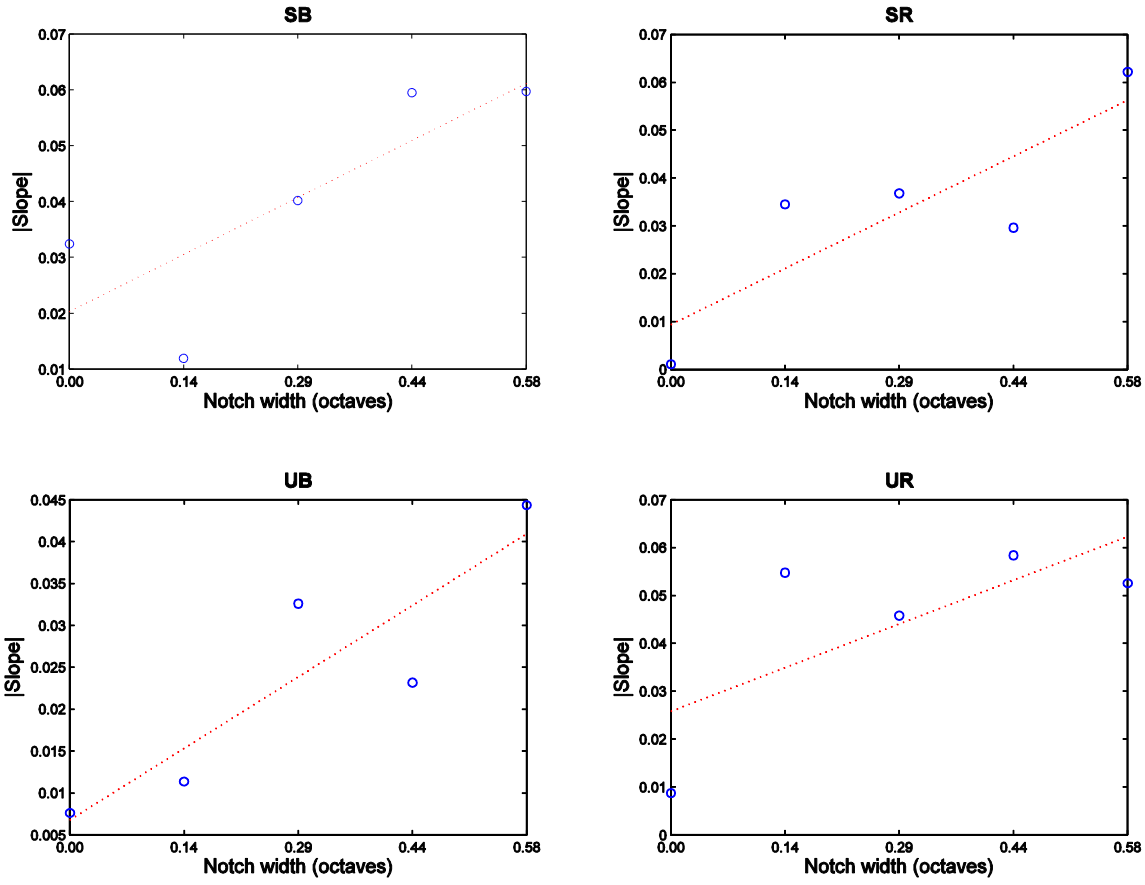

Figure 38: The plots corresponding to each of the four regressions done to find out the "slope of learning" (represented by value b_1 from Table 6 above).

8.4. Discussion

8.4.1 Main analysis

The only significant effect that I found in the hearing experiment was the main effect of notch width – a fact that is, as I have previously mentioned, more than predictable, given the existing literature. There were no other main effects (for the B/R and S/U factors), not interaction effects between nw, B/R and S/U.

This all seems to indicate that the influence of notch width on threshold seems much more robust to context effects in hearing than in vision.

Also particular to the hearing experiment was the fact that the reliability of measurements seemed better: error bars on the hearing graphs (Figures 34-37) are generally smaller than the corresponding graphs from vision (Figures 24-25 and 29-30).

What these two conclusions – robustness and reliability of measurements – suggest is that the dominant mechanism that underlies the relationship between threshold and notch width is more rigid (hard-wired) in hearing than in vision; or, in other words, that thresholds are more deterministic.

8.4.2 Learning effects

The results of the vision experiment indicated that it is under the *blocked* conditions that I am most likely to learn. The situation is similar for hearing, where the regression for learning rate that turned out to be significant was also achieved under a "blocked" (and "unique") condition.

Any further speculation beyond these conclusions would, however, not be scientifically sound, given that, at the moment, the experiment only had data from one observer. Data collection for two more observers is, nevertheless, under way, which will make these conclusions (and perhaps others) equally trustworthy to the ones drawn for the vision experiment.

9. GENERAL DISCUSSION

9.1. Parallels between the vision and the hearing experiments

The fact that I have not found major effects in hearing but I did so in the related experiment – the one in vision – can be interpreted as a revalidation of those effects found in vision. So in a way, I might call the hearing experiment a control experiment for the vision one, even though that is not how I initially thought of them. What I think allows this statement to be true are the important underlying similarities that exist between the two cases – specifically, the channels model – and the similarities between my two experiments. It is important to note, however, that channels in hearing are more likely to originate earlier in the neural processing chain. Evidence for that is, for example, the fact that physiological measurements of the frequency response of nerve fibers in the cochlea are not too dissimilar from psychophysical estimates (Patterson, 1976), as discussed in section 3.6 on auditory filters.

This implies that auditory channels are more rigid than those in vision, a fact that I suggest is also confirmed by the conclusion drawn in section 8.4.1.

In section 4.2.1 I have discussed the plasticity exhibited by the frequency-selective channels in terms of reweighting: it is not these early channels that change when a different stimulus is present, but how they are used at a more central level. In other words, the visual system is "plastic" in the sense of reweighting at a higher level. My research does not directly test whether the flexibility demonstrated in the vision experiment stems from adjustment of low-level channel properties or a higher-level reweighting. However, whatever the mechanism, the effects of this flexibility are evident in the results of a notched-noise masking paradigm.

Just as I have began my thesis by integrating it in the larger picture of existing literature, I should also make a small note regarding a possible implication that my work might have on future research. The past research that I am referring to in the previous sections of the Discussion (and the results of which I termed "safe") is the "traditional" one for this field, i.e. the one concerned primarily with channel estimates measured for normal individuals. However, taking a step back, the techniques for assessing perceptual processing (such as the notched-noise paradigm) have been used in a broader way in

psychology (for example, in the Plaisted et al. paper (2003), which estimated auditory bandwidths in autistic observers). The possibility that context can influence data obtained in notched-noise masking paradigms means that care must be taken in choosing whether to use blocked or randomised conditions, particularly when trying to generalise to, for example, clinical psychology studies of autistic individuals, who may be differently affected by context than normal observers: for example, referring back to Plaisted et al.'s study, their estimates of auditory channel selectivity were significantly different from those obtained previously by Moore (1987) for normal individuals.

In summary, extending the results of this work, which are obtained – like the majority of other research that used notched-noise masking to estimate channel bandwidth – from normal individuals, to clinical cases might not be valid, since the effects of context might be even more critical in such disorders.

9.2. Further work

The same kind of experiments that I have done here in spatial vision (with sinusoidally-varying gratings) could be potentially repeated in temporal vision (with flickering stimuli). There are enough differences and similarities between the two domains to promise a rewarding work of research, that could perhaps lead to such interesting contrasts as have been shown to exist between the pedestal effect in *spatial* vision and in *temporal* vision (described in detail in section 3.4.3).

The search for ways in which mechanisms that were traditionally thought to be hard-wired (such as the channels in vision and hearing) can exhibit plasticity is not limited to the "blocked" vs "randomised" disctinction – other ways in which higher levels can be "baited" to come into play and exhibit perceptual learning can be imagined and tested.

10. APPENDIX A

This appendix shows the order in which the sessions of the vision experiment were run for each individual observer:

Table 7: Outline of data collection schedule for vision experiment.

Sitting no.	Session no.	S/U	B/R	Staircases' notch widths (octaves)
1	1	S	B	0.0, 0.0, 0.0, 0.0, 0.0
	2			0.5, 0.5, 0.5, 0.5, 0.5
	3			1.0, 1.0, 1.0, 1.0, 1.0
	4			1.5, 1.5, 1.5, 1.5, 1.5
	5			2.0, 2.0, 2.0, 2.0, 2.0
2	6	S	R	0.0, 0.5, 1.0, 1.5, 2.0
	7			0.0, 0.5, 1.0, 1.5, 2.0
	8			0.0, 0.5, 1.0, 1.5, 2.0
	9			0.0, 0.5, 1.0, 1.5, 2.0
	10			0.0, 0.5, 1.0, 1.5, 2.0
3	11	U	R	0.0, 0.5, 1.0, 1.5, 2.0
	12			0.0, 0.5, 1.0, 1.5, 2.0
	13			0.0, 0.5, 1.0, 1.5, 2.0
	14			0.0, 0.5, 1.0, 1.5, 2.0
	15			0.0, 0.5, 1.0, 1.5, 2.0
4	16	U	B	0.0, 0.0, 0.0, 0.0, 0.0
	17			0.5, 0.5, 0.5, 0.5, 0.5
	18			1.0, 1.0, 1.0, 1.0, 1.0

Sitting no.	Session no.	S/U	B/R	Staircases' notch widths (octaves)
	19			1.5, 1.5, 1.5, 1.5, 1.5
	20			2.0, 2.0, 2.0, 2.0, 2.0
5	21	S	R	0.0, 0.5, 1.0, 1.5, 2.0
	22			0.0, 0.5, 1.0, 1.5, 2.0
	23			0.0, 0.5, 1.0, 1.5, 2.0
	24			0.0, 0.5, 1.0, 1.5, 2.0
	25			0.0, 0.5, 1.0, 1.5, 2.0
6	26	S	B	0.5, 0.5, 0.5, 0.5, 0.5
	27			1.0, 1.0, 1.0, 1.0, 1.0
	28			1.5, 1.5, 1.5, 1.5, 1.5
	29			2.0, 2.0, 2.0, 2.0, 2.0
	30			0.0, 0.0, 0.0, 0.0, 0.0
7	31	U	B	0.5, 0.5, 0.5, 0.5, 0.5
	32			1.0, 1.0, 1.0, 1.0, 1.0
	33			1.5, 1.5, 1.5, 1.5, 1.5
	34			2.0, 2.0, 2.0, 2.0, 2.0
	35			0.0, 0.0, 0.0, 0.0, 0.0
8	36	U	R	0.0, 0.5, 1.0, 1.5, 2.0
	37			0.0, 0.5, 1.0, 1.5, 2.0
	38			0.0, 0.5, 1.0, 1.5, 2.0
	39			0.0, 0.5, 1.0, 1.5, 2.0
	40			0.0, 0.5, 1.0, 1.5, 2.0
9	41	S	B	1.0, 1.0, 1.0, 1.0, 1.0
	42			1.5, 1.5, 1.5, 1.5, 1.5
	43			2.0, 2.0, 2.0, 2.0, 2.0
	44			0.0, 0.0, 0.0, 0.0, 0.0
	45			0.5, 0.5, 0.5, 0.5, 0.5
10	46	S	R	0.0, 0.5, 1.0, 1.5, 2.0
	47			0.0, 0.5, 1.0, 1.5, 2.0
	48			0.0, 0.5, 1.0, 1.5, 2.0
	49			0.0, 0.5, 1.0, 1.5, 2.0
	50			0.0, 0.5, 1.0, 1.5, 2.0
11	51	U	R	0.0, 0.5, 1.0, 1.5, 2.0

Sitting no.	Session no.	S/U	B/R	Staircases' notch widths (octaves)
	52			0.0, 0.5, 1.0, 1.5, 2.0
	53			0.0, 0.5, 1.0, 1.5, 2.0
	54			0.0, 0.5, 1.0, 1.5, 2.0
	55			0.0, 0.5, 1.0, 1.5, 2.0
12	56	U	B	1.0, 1.0, 1.0, 1.0, 1.0
	57			1.5, 1.5, 1.5, 1.5, 1.5
	58			2.0, 2.0, 2.0, 2.0, 2.0
	59			0.0, 0.0, 0.0, 0.0, 0.0
	60			0.5, 0.5, 0.5, 0.5, 0.5
13	61	S	R	0.0, 0.5, 1.0, 1.5, 2.0
	62			0.0, 0.5, 1.0, 1.5, 2.0
	63			0.0, 0.5, 1.0, 1.5, 2.0
	64			0.0, 0.5, 1.0, 1.5, 2.0
	65			0.0, 0.5, 1.0, 1.5, 2.0
14	66	S	B	1.5, 1.5, 1.5, 1.5, 1.5
	67			2.0, 2.0, 2.0, 2.0, 2.0
	68			0.0, 0.0, 0.0, 0.0, 0.0
	69			0.5, 0.5, 0.5, 0.5, 0.5
	70			1.0, 1.0, 1.0, 1.0, 1.0
15	71	U	B	1.5, 1.5, 1.5, 1.5, 1.5
	72			2.0, 2.0, 2.0, 2.0, 2.0
	73			0.0, 0.0, 0.0, 0.0, 0.0
	74			0.5, 0.5, 0.5, 0.5, 0.5
	75			1.0, 1.0, 1.0, 1.0, 1.0
16	76	U	R	0.0, 0.5, 1.0, 1.5, 2.0
	77			0.0, 0.5, 1.0, 1.5, 2.0
	78			0.0, 0.5, 1.0, 1.5, 2.0
	79			0.0, 0.5, 1.0, 1.5, 2.0
	80			0.0, 0.5, 1.0, 1.5, 2.0
17	81	S	B	2.0, 2.0, 2.0, 2.0, 2.0
	82			0.0, 0.0, 0.0, 0.0, 0.0
	83			0.5, 0.5, 0.5, 0.5, 0.5
	84			1.0, 1.0, 1.0, 1.0, 1.0

Sitting no.	Session no.	S/U	B/R	Staircases' notch widths (octaves)
	85			1.5, 1.5, 1.5, 1.5, 1.5
18	86	S	R	0.0, 0.5, 1.0, 1.5, 2.0
	87			0.0, 0.5, 1.0, 1.5, 2.0
	88			0.0, 0.5, 1.0, 1.5, 2.0
	89			0.0, 0.5, 1.0, 1.5, 2.0
	90			0.0, 0.5, 1.0, 1.5, 2.0
19	91	U	R	0.0, 0.5, 1.0, 1.5, 2.0
	92			0.0, 0.5, 1.0, 1.5, 2.0
	93			0.0, 0.5, 1.0, 1.5, 2.0
	94			0.0, 0.5, 1.0, 1.5, 2.0
	95			0.0, 0.5, 1.0, 1.5, 2.0
20	96	U	B	2.0, 2.0, 2.0, 2.0, 2.0
	97			0.0, 0.0, 0.0, 0.0, 0.0
	98			0.5, 0.5, 0.5, 0.5, 0.5
	99			1.0, 1.0, 1.0, 1.0, 1.0
	100			1.5, 1.5, 1.5, 1.5, 1.5

The order of sessions was the same in the hearing experiment, with the difference that the values of the notch widths were not 0.0, 0.5, 1.0, 1.5 and 2.0 octaves but 0.0000, 0.1444, 0.2895, 0.4361 and 0.5850 octaves (see section 8.2.6 which explains in detail the conditions under which this experiment was run).

11. APPENDIX B

This appendix makes a note about how the trials were counter-balanced as part of the design of each experiment. The vision experiment is discussed below but, as previously stated, the same is valid for the hearing experiment if the notch width values are replaced from 0.0, 0.5, 1.0, 1.5 and 2.0 octaves to 0.0000, 0.1444, 0.2895, 0.4361 and 0.5850 octaves.

Even though for each observer I ran sessions that were counter-balanced in terms of notch width, presentation condition and noise condition, there was one minor counter-balancing requirement that I did not meet, relating to the number of intervening sessions between two consecutive estimates. Ideally, this number would be the same for all measurements of notch width, under a specific set of conditions – for example, there should be the same number of intervening sessions between estimate 1 and estimate 2 of the 0.0-octaves notch width under the SB condition as there are between estimate 1 and estimate 2 of the 0.5-octaves notch width under the same condition, and the same goes for the differences between estimates 2 and 3, estimates 3 and 4 and estimates 4 and 5.

Based on the information in Table 7 (Appendix A), the following table (Table 8) can be derived, which gives an intervening session count between any pair of consecutive estimates, for all notch widths and under all conditions:

Table 8: Number of intervening sessions between consecutive estimates.

S/U	*B/R*	*nw*	*number of intervening sessions between estimates*			
			#1 and #2	*#2 and #3*	*#3 and #4*	*#4 and #5*
S	B	0.0 oct	28	13	23	13
		0.5 oct	23	18	23	13
		1.0 oct	23	13	28	13
		1.5 oct	23	13	23	18
		2.0 oct	23	13	23	13
S	R	0.0 oct	10	20	10	20
		0.5 oct	10	20	10	20
		1.0 oct	10	20	10	20
		1.5 oct	10	20	10	20

		2.0 oct	10	20	10	20
U	B	0.0 oct	18	23	13	23
		0.5 oct	13	28	13	23
		1.0 oct	13	23	18	23
		1.5 oct	13	23	13	28
		2.0 oct	13	23	13	23
U	R	0.0 oct	20	10	20	10
		0.5 oct	20	10	20	10
		1.0 oct	20	10	20	10
		1.5 oct	20	10	20	10
		2.0 oct	20	10	20	10

The highlighted numbers represent the notch widths that were "unfavoured", in the sense that more sessions have passed between two of their estimates than have passed for the rest of the notch width values under the same condition, between the same two consecutive estimates. It can be seen that it is the 2.0-octaves notch width that is *never* favoured, which is the same notch width for which learning has been previously shown to be manifested the strongest.

It would be rather complicated to try and correct this artefact, and even so, the correction would most likely not alter my results fundamentally. The precision of my results was less than perfect anyway because the time lapse between two consecutive "sittings" (i.e. block of 5 sessions) was not homogenous across observers.

12. ACKNOWLEDGEMENTS

I first and foremost thank my supervisor, Hannah Smithson, who patiently bore with me through interminable lists of questions, and who opened the door of Psychology to me.

I also thank my colleague Wayne Smith, who often helped me along the difficult road that leads into Psychophysics, and also provided valuable feedback before my public talks.

My colleague Cristiana has been a great friend throughout my time in Durham and has offered me invaluable moral support, as well as suggestions for improving my thesis and relating to the work of research in general, and for this I give her a special thank you.

Last but not least, I thank my mother for offering me the chance to make an unexpected (or maybe not so unexpected) shift of career from Engineering to Psychology, a step that I hope this MRes will only be the first step of.

13. REFERENCES

Adini, Y., Wilkonsky, A., Haspel, R., Tsodyks, M., & Sagi, D. (2004). Perceptual learning in contrast discrimination: The effect of contrast uncertainty. *Journal of Vision, 4*(12), 993-1005.

Ahumada, A., & Beard, B. (1997). Image discrimination models predict detection in fixed but not random noise. *Journal of the Optical Society of America A, 14*(9), 2471-2476.

Ahumada, A., & Beard, B. (1997). *Image discrimination models: Detection in fixed and random noise.* Paper presented at the SPIE.

Albrecht, D., & De Valois, R. (1981). Striate cortex responses to periodic patterns with and without the fundamental harmonics. *The Journal of Physiology, 319*(1), 497-514.

Baker, R., & Rosen, S. (2006). Auditory filter nonlinearity across frequency using simultaneous notched-noise masking. *Journal of the Acoustical Society of America, 119*(1), 454-462.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*(8), 617-629.

Blakemore, C., & Campbell, F. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology, 203*(1), 237-260.

Braddick, O., Campbell, F., & Atkinson, J. (1978). Channels in vision: Basic aspects. *Handbook of Sensory Physiology, 8*, 3-38.

Burgess, A. (1985). Visual signal detection. III. On Bayesian Use of Prior Knowledge and Cross-Correlation. *Journal of the Optical Society of America A, 2*(9), 1498-1507.

Burgess, A., & Ghandeharian, H. (1984). Visual signal detection. I. Phase sensitive detection. *Journal of the Optical Society of America A, 1*, 900-905.

Burgess, A., & Ghandeharian, H. (1984). Visual signal detection. II. Signal-location identification. *Journal of the Optical Society of America A, 1*, 906-910.

Campbell, D., Stanley, J., Gampbell, D., & Stanley, J. (1969). *Experimental and quasi-experimental designs for research*: Rand McNally Chicago.

Campbell, F. W., Cooper, G. F., & Enroth-Cugell, C. (1969). The spatial selectivity of the visual cells of the cat. *The Journal of Physiology, 203*(1), 223-235.

Campbell, F. W., & Kulikowski, J. (1966). Orientational selectivity of the human visual system. *The Journal of Physiology, 187*(2), 437-445.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology, 197*(3), 551-566.

Cornsweet, T., & Pinsker, H. (1965). Luminance discrimination of brief flashes under various conditions of adaptation. *The Journal of Physiology, 176*(2), 294-310.

de Boer, E. (1967). Correlation studies applied to the frequency resolution of the cochlea. *Journal of Auditory Research, 7*, 209-217.

De Valois, R. L., & De Valois, K. K. (1980). Spatial vision. *Annual Review of Psychology, 31*, 309-341.

Derrington, A., & Henning, G. (1989). Some observations on the masking effects of two-dimensional stimuli. *Vision Research, 29*(2), 241.

Dosher, B., & Lu, Z. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting (Vol. 95, pp. 13988-13993): National Acad Sciences.

Dosher, B., & Lu, Z. (1999). Mechanisms of perceptual learning. *Vision Research, 39*(19), 3197-3221.

Dwyer, D. (2008). Perceptual Learning: Complete Transfer across Retinal Locations. *Current Biology, 18*(24), 1134-1136.

Eckstein, M., Ahumada, A. J., & Watson, A. (1997). Image discrimination models predict human detection in medical images. *Human Vision, Visual Processing, and Digital Display VIII, 3016*, 44-56.

Eckstein, M., Ahumada, A. J., & Watson, A. (1997). Visual signal detection in structured backgrounds. II. Effects of contrast gain control, background variations, and white noise. *Journal of the Optical Society of America A, 14*(9), 2406-2419.

Foley, J., & Legge, G. (1981). Contrast detection and near-threshold discrimination in human vision. *Vision Research, 21*(7), 1041-1053.

Glasberg, B., & Moore, B. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research, 47*(1-2), 103.

Grossberrg, S. (1974). Classical and instrumental learning by neural networks. *Progress in theoretical biology, 3*(51-141), 42-47.

Henning, G. B. (1969). Amplitude discrimination in noise, pedestal experiments, and additivity of masking. *Journal of the Acoustical Society of America, 45*(1), 327-328.

Henning, G. B., Hertz, B., & Broadbent, D. (1975). Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. *Vision Research, 15*, 887–897.

Henning, G. B., & Wichmann, F. A. (2007). Some observations on the pedestal effect. *Journal of Vision, 7*(1), 1-15.

Houtgast, T. (1974). *Lateral suppression in hearing.* Free University, Amsterdam.

Hurlbert, A. (2000). Visual perception: Learning to see through noise. *Current Biology, 10*(6), 231-233.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences, 88*(11), 4966-4970.

Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature, 365*(6443), 250-252.

Keppel, G., & Wickens, T. (1982). *Design and analysis: A researcher's handbook*: Prentice-Hall Englewood Cliffs, NJ.

Knill, D., Friedman, W., & Geisler, W. (2003). Bayesian and Statistical Approaches to Vision. *Journal of the Optical Society of America A, 20*(7), 1232-1233.

Kulikowski, J., & Robson, A. (1999). Spatial, temporal and chromatic channels: electrophysiological foundations. *Journal of Optical Technology, 66*(9), 797-806.

Legge, G., & Foley, J. (1980). Contrast masking in human vision. *Journal of the Optical Society of America A, 70*(12), 1458-1471.

Legge, G., Kersten, D., & Burgess, A. (1987). Contrast discrimination in noise. *Journal of the Optical Society of America A, 4*(2), 391-404.

Leshowitz, B., Taub, H., & Raab, D. (1968). Visual detection of signals in the presence of continuous and pulsed backgrounds. *Percept. Psychophys, 4*, 207–213.

Liu, Z., & Vaina, L. (1998). Simultaneous learning of motion discrimination in two directions. *Cognitive Brain Research, 6*(4), 347-349.

Losada, M., & Mullen, K. (1995). Color and luminance spatial tuning estimated by noise masking in the absence of off-frequency looking. *Journal of the Optical Society of America A, 12*(2), 250-260.

Lu, Z., & Dosher, B. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *Journal of the Optical Society of America A, 16*(3), 764-778.

Lu, Z., & Dosher, B. (2001). Characterizing the spatial-frequency sensitivity of perceptual templates. *Journal of the Optical Society of America A, 18*(9), 2041-2053.

Margolis, R., & Small, A. (1975). The Measurement of Critical Masking Bands. *Journal of Speech, Language and Hearing Research, 18*(3), 571-587.

Marr, D. (1983). *Vision: a computational investigation into the human representation and processing of visual information*: WH Freeman.

Mollon, J., & Danilova, M. (1996). Three remarks on perceptual learning. *Spatial Vision, 10*(1), 51-58.

Moore, B. (1987). Distribution of auditory-filter bandwidths at 2 kHz in young normal listeners. *Journal of the Acoustical Society of America, 81*(5), 1633-1635.

Moore, B., & Glasberg, B. (1981). Auditory filter shapes derived in simultaneous and forward masking. *Journal of the Acoustical Society of America, 70*(4), 1003-1014.

Nachmias, J., & Sansbury, R. (1974). Letter: Grating contrast: discrimination may be better than detection. *Vision Research, 14*(10), 1039-1042.

Patterson, R. D. (1974). Auditory filter shape. *Journal of the Acoustical Society of America, 55*(4), 802-809.

Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. *Journal of the Acoustical Society of America, 59*(3), 640-654.

Patterson, R. D., & Henning, G. B. (1977). Stimulus Variability and Auditory Filter Shape. *Journal of the Acoustical Society of America, 62*(3), 649-664.

Peli, E., Garcia-Perez, M. A., Giorgi, R., & Woods, R. L. (2004). Spatial or temporal 2AFC may give different results depending on context. *Journal of Vision, 4*(8), 779-779.

Pelli, D. G. (1981). *Effects of visual noise.* University of Cambridge.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A, 2*(9), 1508-1531.

Petrov, A., Dosher, B., & Lu, Z. (2005). The dynamics of perceptual learning: An incremental reweighting model. *Psychological Review, 112*(4), 715-743.

Pfafflin, S., & Mathews, M. (1961). An Energy Detection Model for Monaural Auditory Detection. *Journal of the Acoustical Society of America, 33*(11), 1654-1655.

Plaisted, K., Saksida, L., Alcantara, J., & Weisblatt, E. (2003). Towards an understanding of the mechanisms of weak central coherence effects: experiments in visual configural learning and auditory perception. *Phil. Trans. R. Soc. B, 358*(1430), 375-386.

Raab, D., Osman, E., & Rich, E. (1963). Intensity Discrimination, the "Pedestal" Effect, and "Negative Masking" with White-Noise Stimuli. *Journal of the Acoustical Society of America, 35*(7), 1053-1053.

Robson, J. (1975). Receptive fields: Neural representation of the spatial and intensive attributes of the visual image. *Handbook of Perception, 5*, 81–116.

Rumelhart, D., & McClelland, J. (1986). the PDP Research Group, Parallel Distributed Processing: Explorations in the Microstructures of Cognition: Cambridge, MA: MIT Press.

Sachs, M. B., Nachmias, J., & Robson, J. G. (1971). Spatial-Frequency Channels in Human Vision. *Journal of the Optical Society of America A, 61*(9), 1176-1186.

Schade Sr, O. (1956). Optical and photoelectric analog of the eye. *Journal of the Optical Society of America A, 46*(9), 721-739.

Sekuler, R. (1974). Spatial vision. *Annual Review of Psychology, 25*, 195-232.

Smallman, H., MacLeod, D., He, S., & Kentridge, R. (1996). Fine grain of the neural representation of human spatial vision. *Journal of Neuroscience, 16*(5), 1852-1859.

Smithson, H., Henning, B., MacLeod, D., & Stockman, A. (2009, under review). The effect of notched noise on flicker detection and discrimination. *Journal of Vision*.

Solomon, J., & Pelli, D. (1994). The visual filter mediating letter identification. *Nature, 369*(6479), 395-397.

Stromeyer, C. F., III, & Julesz, B. (1972). Spatial-Frequency Masking in Vision: Critical Bands and Spread of Masking. *Journal of the Optical Society of America A, 62*(10), 1221-1232.

Swift, D., & Smith, R. (1983). Spatial frequency masking and Weber's Law. *Vision Research, 23*(5), 495-505.

Taylor, C. P., Bennett, P. J., & Sekuler, A. B. (2006). Narrow-band channels optimally sum a broad band of spatial frequency information [Abstract]. *Journal of Vision, 6*(6), 115-115.

Tolhurst, D., Tadmor, Y., & Chao, T. (1992). Amplitude spectra of natural images. *Ophthalmic and Physiological Optics, 12*(2), 229-232.

Watson, A. B., Borthwick, R., & Taylor, M. (1997). *Image quality and entropy masking*.

Williams, E. (2004). Critical Bands in Vision. University of Cambridge.

Zwicker, E., Flottorp, G., & Stevens, S. (1957). Critical Band Width in Loudness Summation. *Journal of the Acoustical Society of America, 29*(5), 548-557.