– 2009 –

KATHOLIEKE UNIVERSITEIT LEUVEN

Faculteit Psychologie en Pedagogische Wetenschappen

Laboratorium voor Experimentele Psychologie

Early Stages in Spatial Vision – Psychophysical Data and Computational Models of Contrast Perception

Proefschrift aangeboden

tot het verkrijgen van de graad van

Doctor in de Psychologie

door

Robbe L. T. Goris

Promotor: Prof. Dr. Johan Wagemans – Co-promotor: Prof. Dr. Felix Wichmann

Robbe L. T. Goris

Early stages in spatial vision – psychophysical data and computational models of contrast perception

Dissertation submitted to obtain the degree of Doctor in Psychology, March 2009
Supervisor: Prof. Dr. Johan Wagemans Co-supervisor: Prof. Dr. Felix Wichmann

All current models of spatial vision postulate an initial linear spatial frequency- and orientation-selective filtering stage. Because sinusoids are the eigenfunctions of any linear system, sinusoidal gratings have become perhaps the single most popular stimuli to investigate early vision. Often, visual noise is added to these stimuli in behavioural experiments designed to characterize certain aspects of the visual mechanisms mediating pattern vision. External noise methods are popular because noise is a random stimulus – and thus able to yield unbiased results –, yet its characteristics are highly controllable. However, effects of visual noise on the processing of sinusoidal gratings are not well understood. At the neurophysiological level, there have been few attempts to systematically explore contrast processing in the presence of visual noise. At the psychophysical level, assumptions regarding noise effects vary from excitation to inhibition. The central goal of this thesis is to advance our understanding of the visual processing of sinusoidal gratings embedded in visual noise. To this end, behavioural measurements of sinusoidal contrast detection and discrimination in noise are conducted and computational models of contrast perception in spectrally flat and filtered noise are developed.

The most prominent computational models developed to account for contrast detection in noise *all* predict linear threshold-vs.-noise-contrast functions. Here, it is argued that this prediction is inconsistent with the standard model of the early visual system as derived from sinusoidal contrast discrimination experiments and neurophysiological findings. In the standard model, the linear filtering stage is followed by a (rectifying) nonlinearity and internal noise. Such linear-non-linear cascade models predict an improvement in observers' contrast detection performance when low, sub-threshold levels of external noise are added (i.e., stochastic resonance). Carefully designed behavioural measurements of contrast detection in noise were conducted in two experiments. Detection thresholds were found to reach a minimum for very low noise levels. It is demonstrated that an elaborated version of the divisive contrast-gain control model fits these data reasonably well.

In the gain-control model, a single mechanism underlies stochastic resonance and the pedestal effect – the pedestal effect is the improvement in the detectability of a sinusoidal grating in the presence of another low-contrast grating of the same orientation, spatial frequency and phase. An alternative hypothesis may be to argue that both effects characterize two independent mechanisms. To address this issue, sinusoidal contrast discrimination in weak and modest levels of noise was measured in a third experiment. Making use of a full quantitative description of the data combined with comprehensive model selection assessments, the pedestal effect is shown to be more reduced in the presence of weak noise than in moderate noise. This reduction rules out independent, additive sources of performance improvement and supports the parsimonious explanation that a single mechanism underlies the pedestal effect and stochastic resonance in contrast perception. A simulation study demonstrates that the elaborated gain-control model can produce similar results.

Recent evidence has demonstrated that the pedestal effect in spatial vision is differently modified by spectrally flat and notch-filtered noise: the pedestal effect is reduced in flat noise, but virtually disappears in the presence of notched noise (Henning & Wichmann, 2007). As all other current models of spatial vision, the standard gain-control model fails to account for these results. A new population-code model for human contrast discrimination is proposed. It is demonstrated that combining the outputs of many units whose contrast response functions resemble those of the cortical cells believed to underlie human pattern vision by simple response-based pooling, produces contrast-discrimination data consistent with psychophysical observations: the pedestal effect is present without noise, reduced in broadband noise, but almost disappears in notched noise. These findings follow naturally from the normalization model of simple cells in primary visual cortex followed by response-based pooling, and suggest that in processing even low-contrast sinusoidal gratings, the visual system may combine information across neurons tuned to different spatial frequencies and orientations.

Robbe L. T. Goris

Vroege stadia in patroonperceptie – psychofysische data en computationele modellen

Proefschrift aangeboden tot het verkrijgen van de graad van Doctor in de Psychologie, maart 2009
Promotor: Prof. Dr. Johan Wagemans Co-promotor: Prof. Dr. Felix Wichmann

Alle huidige visuele perceptie modellen stellen dat het eerste visuele informatieverwerkingsstadium bestaat uit lineaire, spatiale frequentie- en oriëntatiegevoelige filtering. Omdat sinusfuncties de eigenfuncties zijn van ieder lineair systeem behoren sinusgolfpatronen tot de populairste stimuli om vroege visuele processen te onderzoeken. In gedragsexperimenten ontworpen om specifieke aspecten van visuele informatieverwerking te onderzoeken wordt vaak ruis toegevoegd aan deze stimuli. Methodes die gebruik maken van externe ruis zijn populair omdat ruis enerzijds een random stimulus is – en dus niet-gebiaste resultaten oplevert –, maar anderzijds erg manipuleerbaar is. Toch is het begrip van de effecten van visuele ruis op de verwerking van sinusgolfpatronen beperkt. Slechts weinig neurofysiologische studies hebben contrastverwerking in de aanwezigheid van ruis systematisch onderzocht. In de psychofysische literatuur variëren de assumpties aangaande ruiseffecten van excitatie tot inhibitie. De belangrijkste doelstelling van deze thesis is om ons begrip van de visuele verwerking van sinusgolfpatronen in ruis te verbeteren. Hiervoor worden gedragsmetingen van contrastdetectie en –discriminatie in ruis uitgevoerd en computationele modellen van contrastperceptie in spectraal vlakke en gefilterde ruis ontwikkeld.

De belangrijkste mathematische modellen voor contrast detectie in ruis voorspellen *allemaal* lineaire threshold-vs-ruiscontrast functies. In dit proefschrift wordt beargumenteerd dat deze predictie inconsistent is met het standaardmodel voor het vroege visuele systeem dat afgeleid is uit contrastdiscriminatie experimenten en neurofysiologische bevindingen. In het standaardmodel wordt lineaire filtering gevolgd door een (rectificerende) niet-lineariteit en interne ruis. Dergelijke lineaire-niet-lineaire cascade modellen voorspellen een verbetering van de contrastdetectieperformantie in de aanwezigheid van zwakke, sub-threshold ruis (i.e., stochastische resonantie). Zorgvuldig ontworpen gedragsmetingen van contrastdetectie in ruis werden uitgevoerd in twee experimenten. De resultaten laten zien dat detectiethresholds een minimum bereiken bij erg lage ruisniveaus. Een aangepaste versie van het contrast-gain control model beschrijft deze data goed.

In het gain-control model veroorzaakt eenzelfde mechanisme stochastische resonantie en het pedestal effect – het pedestal effect is de verbetering in de detectabiliteit van een sinusgolf-patroon in aanwezigheid van een ander laag-contrast patroon met dezelfde oriëntatie, spatiale frequentie en fase. Een alternatieve interpretatie is dat beide effecten de werking van twee onafhankelijke mechanismes karakteriseren. Om deze hypothese te testen werd contrastdiscriminatie in zwakke en sterke ruis gemeten in een derde experiment. Gebruik makend van een volledige kwantitatieve beschrijving van de data en een uitgebreide modelselectieprocedure wordt aangetoond dat het pedestal effect sterker afneemt in zwakke ruis dan in sterke ruis. Deze afname sluit onafhankelijke, additieve bronnen van performantieverbetering uit en ondersteunt de spaarzame verklaring dat eenzelfde mechanisme ten grondslag ligt aan het pedestal effect en stochastische resonantie in contrastperceptie. Een simulatiestudie toont aan dat het contrast-gain control model gelijkaardige resultaten kan produceren.

Recente evidentie heeft aangetoond dat het pedestal effect in contrastperceptie verschillend beïnvloed wordt door spectraal vlakke en notch-gefilterde ruis: het pedestal effect neemt af in vlakke ruis, maar verdwijnt bijna volledig in de aanwezigheid van notch-gefilterde ruis (Henning & Wichmann, 2007). Zoals alle andere huidige contrastperceptiemodellen kan het gain-control model deze resultaten niet verklaren. Een nieuw populatiecoderingsmodel voor contrastdiscriminatie wordt voorgesteld. Er wordt aangetoond dat het combineren van de output van verschillende units met contrastresponsfuncties gelijkend op de corticale cellen onderliggend aan patroonperceptie op basis van responsgebaseerde weging contrastdiscriminatiedata produceert die consistent zijn met de psychofysische observaties: het pedestal effect is aanwezig zonder ruis, gereduceerd in witte ruis en nagenoeg verdwenen in notch-gefilterde ruis. Deze bevindingen suggereren dat zelfs in de verwerking van laag-contrast sinusgolfpatronen, het visueel systeem informatie combineert over neuronen getuned op verschillende spatiale frequenties en oriëntaties.

Acknowledgements

First and foremost, I wish to thank my both supervisors, Johan Wagemans and Felix Wichmann. Johan initiated my interest in visual neuroscience and provided continuous support during my stay in his lab. He encouraged me to explore research areas outside the lab's tradition, and I am grateful for his support and encouragement. Felix taught me many of the technical and mathematical skills required to measure and model psychophysical data. Visiting his lab in Tübingen and later on in Berlin was always very rewarding and enjoyable. Much of my scientific inspiration originated in his work and in our conversations on a variety of themes.

Many thanks go to Peter Zaenen for introducing me to psychophysics and insisting to conduct only proper scientific work. I also wish to express my sincere gratitude to Bruce Henning for providing so much unofficial supervision and priceless advice. I thank Tom Putzeys for sharing my passion for early vision and psychophysics and making my stay in the lab much more enjoyable than it would have been without him.

Hans Op de Beeck, Martin lages, Michael Shadlen, Pedro Rosas and Frank Jäkel all deserve a special mentioning for providing inspiration, support and crucial insights. Ben Schouten, Hannah Dold, Ellen Gillard, Bart Machilsen, Bart Ons, Lizzy Bleumers, Geir Nygard, Sven Panis, Marc Van Baelen, Joachim Vandekerckhove and many other colleagues all contributed to my daily happiness at work, for which I express my gratitude.

In 2007, I was allowed to be a student at the Cold Spring Harbor workshop on visual neuroscience, which was scientifically rewarding and inspiring. I thank the organizers for this opportunity. Financial support was provided through a personal grant obtained from the FWO – Vlaanderen and I am grateful for this support.

Most of all, I am grateful to Els Eerdekens and Lukas Goris for sharing their lifes with me.

Contents

Chapter 1 General introduction

This chapter provides a general background for the topics addressed in this thesis, as well as a motivation for the approach used throughout the following chapters. Some of the historical roots, main findings and issues of debate in the field of spatial vision are discussed briefly. Further, the main achievements and limitations of current theories and models of the psychophysical and neurophysiological mechanisms underlying contrast perception are addressed briefly. Finally, the link between this general framework and the central issues investigated in the following chapters is clarified.

1.1 Early stages in spatial vision – psychophysical data and computational models of contrast perception

1.1.1 Why study spatial vision?

All animal interaction with the outside world ultimately depends on sensory information. Senses are thus critical for survival. Of all human senses, the visual system is the most complex – approximately one third of the human brain is engaged in vision. Starting from electromagnetic waves hitting the retina, our visual system manages to infer what is where in the environment. Because this process is remarkably fast and accurate, we hardly ever consider the fundamental problems overcome by the visual system.

The retinal image is not only determined by the relevant 'world information', but also by many usually irrelevant factors such as the level of light, the location of the light source(s), the position of the observer in the scene, the spatial relation between objects in the scene, etc. There is thus no simple correspondence between the information in the visual stimulus and the interesting aspects of the world that is being perceived. Therefore, each visual system roughly consists of two stages: a first encoding stage in which the raw input image is broken down into fundamental parts in several steps, and a second decoding stage in which these parts are used to construct visual percepts (and thus to compute what is where). Presumably, it is easier to infer what is where from the fundamental parts than from the raw input image (Graham, 1989).

Stated quite broadly, spatial vision research aims to understand early visual processing of spatial information, i.e., to understand how the neurophysiological and psychophysical mechanisms that allow us to represent and identify spatial patterns work. In general, current visual science has a much better understanding of the encoding stage than of the decoding stage, and most of what is known, concerns the first steps of the encoding process.

1.1.2 Why measure psychophysical data?

Psychophysics is, in short, the scientific study of the relation between stimulus and sensation and deals with *the analysis of perceptual processes by studying the effect on a subject's experience or behaviour of systematically varying the properties of a stimulus along one or more physical dimensions* (Bruce, Green & Georgeson, 1996). The origins of many techniques and theories of current psychophysics go back to the mid-nineteenth century, when 'founder' Gustav Theodor Fechner published *Elemente der Psychophysik* (Fechner, 1860). Similarly important was the development of *Signal Detection Theory* (SDT), approximately half a century ago (Green & Swets, 1966; Macmillan & Creelman, 1991; Tanner & Swets, 1954; Wickens, 2002). SDT provides a general theoretical framework to consider information processing and decision making in uncertain situations. Making use of a limited number of assumptions and a firm mathematical basis, SDT can elegantly be applied to the design and analysis of behavioural experiments that investigate sensory processing.

Consider a person who has to decide whether or not some specific condition is present, using information that is incomplete, ambiguous and often contradictory. Assuming that (1) the evidence extracted by this person – representing (something akin to) the likelihood of this condition – can be represented on a single, continuous dimension, i.e., the decision statistic; (2) the extracted evidence is subject to random variation and (3) the person's decision is determined by applying a simple decision criterion to the magnitude of the evidence, it is possible to separate characteristics of the person's decision statistic – sometimes referred to as internal stimulus representation – from the person's response strategies. Founded on no more than these three assumptions, SDT thus explains how choice responses ("condition present" or "condition absent") are related to an observer's inherent ability to discriminate between two physical situations (the presence or absence of the condition).

These insights have been widely applied in psychophysics, as they allow a characterization of an observer's decision statistic as a function of physical stimulus manipulations. Thus, choice responses may provide much information on an observer's internal stimulus representation.

1.1.3 Why use computational models?

Computational models aim to give a compact description, in a mathematical language, of the relation between stimulus and/or observer characteristics and behavioural task performance in perceptual tasks. Formulating such relations is in a sense at the heart of psychophysics, as the field was established by the first successful attempts to relate stimulus to perception using functional equations as mathematical models (Fechner, 1860).

In spatial vision, both the experimental tasks and computational models fit well in the general framework provided by SDT (Green & Swets, 1966; Nachmias, 1972). More specifically, choice responses are assumed to directly reflect the characteristics of the internal response distributions or decision statistic used by the observers. In most empirical SDT-applications, Gaussian internal response distributions prove adequate because the decision statistic is only determined up to a monotonic transformation. For most reasonably behaved decision statistics, there will be a monotonic transformation that makes the decision statistic Gaussian. Consequently, because Gaussian distributions are fully determined by the mean and variance, computational models suffice by specifying these two parameters of the decision statistic for each stimulus condition. Typically, the mean and variance of the decision statistic at the overall system level depend on several component processes within the model. These component processes, often derived from sensory psychology and physiology, try to capture both inefficiencies in the perceptual system and the computations performed by the perceptual process (e.g., Lu & Dosher, 2008).

In sum, to explain observer performance in basic visual tasks, computational models make use of component processes to specify the mean and variance of decision statistics. Transforming the decision statistics to performance measures allows one to compare the predictions of the models and empirical data quantitatively. For the two-alternative-forced-choice designs used in this thesis, this comparison is ultimately based on the assumption that choice responses are generated by Bernoulli processes, i.e., for any given stimulus condition, both response categories have a certain, fixed probability that remains constant throughout the whole experiment.

1.2 General background

1.2.1 The fundamentals: Visual processing near threshold

Our ability to perceive spatial patterns depends both on the contrast and the spatial scale of the visual information (see Figure 1-1). The first measurements of the contrast-sensitivity function (CSF) of the human visual system – i.e., the inverse of the contrast detection threshold as a function of the spatial frequency of single frequency sinusoidal gratings – were conducted by Schade (1956) and interpreted as the human modulation transfer function (MTF).

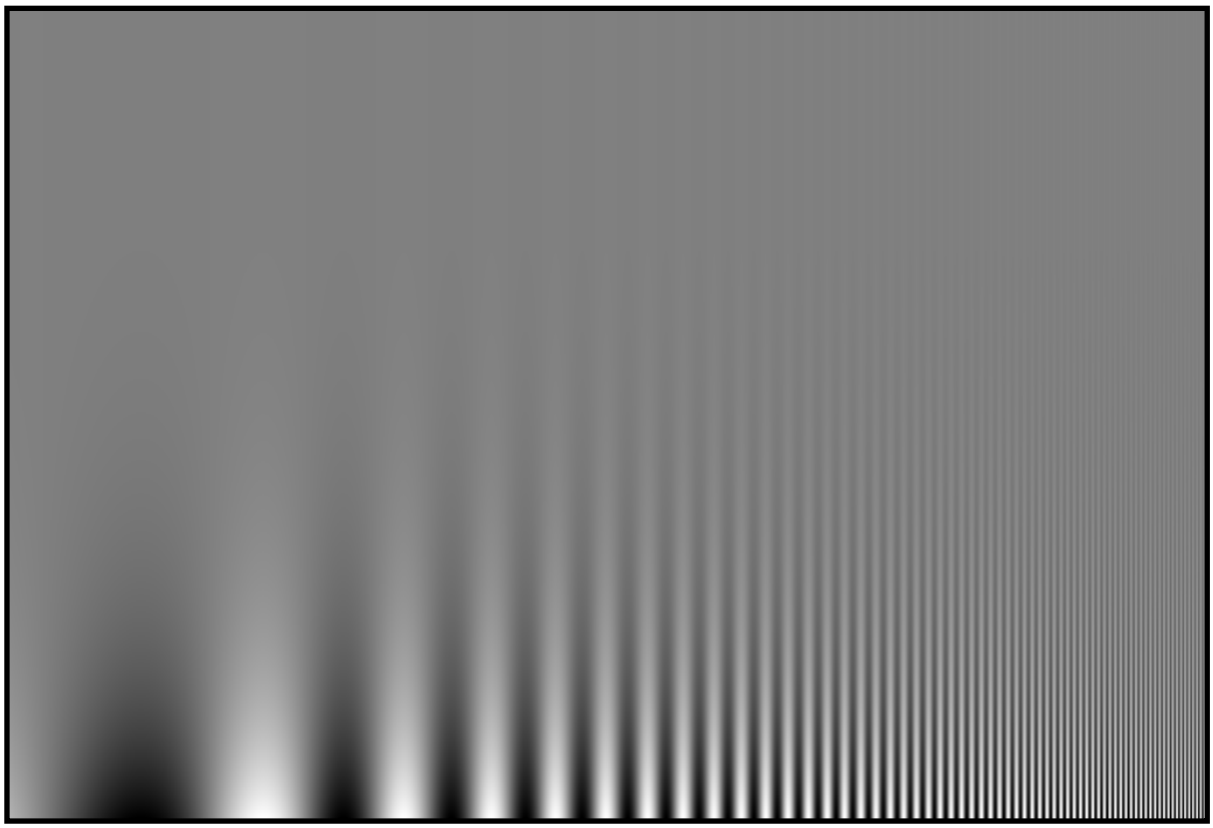

Figure 1-1. *A replication of the Campbell-Robson CSF chart, originally used by Campbell and Robson as a hand-out at the OSA meeting of 1964. Spatial frequency increases logarithmically along the horizontal dimension and contrast decreases logarithmically along the vertical dimension. This test image is well suited to demonstrate the form of the CSF. At a certain height, the alternating light and dark bars can no longer be discriminated from a uniform field of grey (provided that the display-quality of the image is sufficiently high). The higher the bars appear the more sensitive you are for the corresponding spatial frequency. The bars in the middle appear higher than the bars on the side, indicating the inverted U-shape of the CSF. (Aliasing-problems in printing this figure may cause artefacts at the heighest spatial frequencies)*

As can be derived from Figure 1-1, contrast sensitivity may initially increase with spatial frequency, but from approximately 4 cycles per degree (cpd) on, contrast sensitivity decreases with spatial frequency. The attenuated sensitivity at high spatial frequencies is primarily optically driven, i.e., caused by light diffraction and the modulation transfer function of the human eye (Banks, Geisler & Bennett, 1987; Campbell & Green, 1965; Campbell & Gubish, 1966; Schade, 1956). However, the reduced contrast sensitivity at low spatial frequencies cannot be explained in terms of optical limits, but is instead largely driven by the nervous system.

Campbell and Robson (1968) investigated to what extent linear systems analysis forms a useful tool to study contrast sensitivity around detection threshold. In linear systems, superposition holds, meaning that the response caused by the sum of two or more stimuli equals the sum of responses caused by each stimulus individually. Because the Fourier theorem states that any reasonably continuous 2-D wave can be expressed as the sum of a finite number of sine and cosine terms, each with its own phase, spatial frequency, orientation and contrast; responses of a linear visual system to any given stimulus can be predicted from the MTF of that system. Indeed, using the CSF as estimate for the human MTF, Campbell and Robson found that detection thresholds for any given pattern can be predicted from the CSF.

More specifically, around detection threshold, the visual system behaves as if it is composed of a series of independent, spatially located, narrowly tuned, linear, spatial-frequency and orientation-selective filters. Each filter is followed by a response threshold and together, filter and threshold constitute a channel. These findings suggest that the early visual system makes use of *a spatially localised spatial scale code* to process visual information (Georgeson, May, Freeman & Hesse, 2007). Presumably, the visual filters or channels are labelled and signal perceptually relevant information such as the location of large visual objects (i.e., low spatial frequencies), the local orientation of object-parts and -borders (i.e., high spatial frequencies), the velocity of objects, the local image contrast etc. The '*multi-channel model'* can explain a vast majority of near-threshold data and captures many crucial aspects of early spatial vision (Graham, 1989). It is illustrated in Figure 1-2.

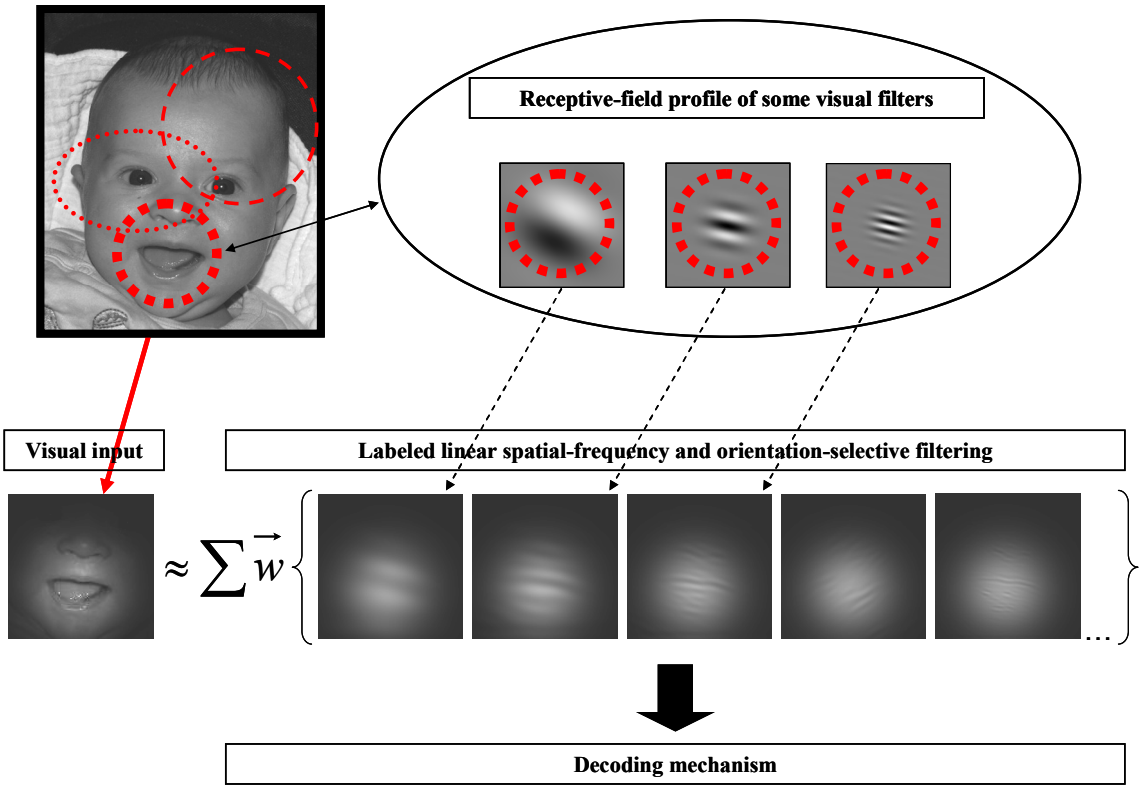

Figure 1-2. *According to the multi-channel model, the early visual system processes visual information in spatially localised spatial frequency- and orientation-selective filters. Above left: the visual stimulus. The red circles and ellipsoids illustrate the spatial location or receptive field for some visual filters. Above right: illustration of the optimal stimuli for three different filters, having the same receptive field, but sensitive for different orientations and spatial frequencies. Below: illustration of the responses of some filters with the same receptive field. Weighted summation of linear filter responses may be able to approximate the input image, provided that the image representation is not overcomplete (see Simoncelli & Freeman, 1995). For the Gabor-style receptive fields illustrated above right, however, the image representation is usually not easily invertible.*

To test the multi-channel model quantitatively, the output of these channels must be connected to the observer's responses in psychophysical experiments, i.e. a decoding mechanism must be specified. It has been well established that a simple maximum-output rule suffices for near-threshold detection studies (Graham, 1989). The maximum output rule states that the response of the observer is based on the comparison of the largest of all channel outputs with some decision criterion. If a blank is presented, any channel may generate the largest output (i.e., internal noise); if a low-contrast signal is presented, the channel most sensitive for the signal is most likely to generate the largest output (Pelli, 1985). Figure 1-3 illustrates these notions.

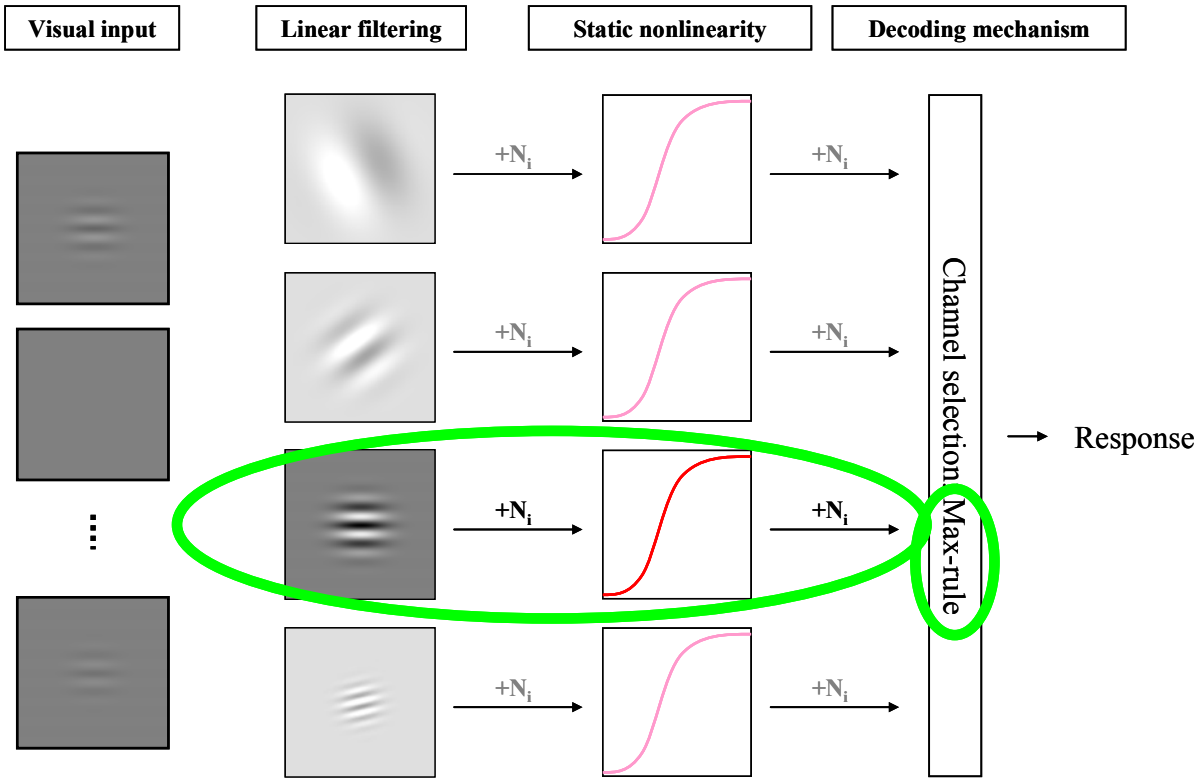

Figure 1-3. *Computational modelling of near-threshold vision. Threshold detection studies – using visual stimuli as illustrated under 'visual input' – have provided many behavioural results that are consistent with a model of our visual system in which linear filtering is followed by a static nonlinearity, internal noise (N_i) and a simple Max-rule as decoding mechanism (this means that the observer's response is based on the channel with maximal output). The Max-rule implies that behavioural measurements with single component sinusoidal gratings are characteristic of the single channel best tuned to the signal.*

On other occasions, no early internal noise is assumed, which makes the Max-rule identical to the assumption that the observer monitors the best-tuned channel (Graham, 1989). Thus, to explain near-threshold vision, higher levels of visual processing need not be modelled. In the words of Norma Graham (1989, pp. 12 and 1992, pp. 57): *"...It is (or we can hope it is) as if the simplicity of the near-threshold experimental situation has made all the higher levels of visual processing transparent, allowing the properties of the low-level multiple analyzers to shine through. Why this should be so is something of a mystery..."*

1.2.2 Visual processing of supra-threshold spatial information

In natural environments, our visual system is confronted with a wide range of contrasts, the majority of which is suprathreshold (for an appropriate analysis of the statistical properties of natural images, see Frazor & Geisler, 2006). Thus, to be useful, any model of spatial vision must aim to explain supra-threshold visual processing. In general, attempts to explain the results of supra-threshold measurements making use of a linear filtering stage and a maximum output rule as decoding mechanism have been unsatisfactory (Graham, 1989). Studies making use of single-component sinusoidal gratings typically find detection-threshold effects indicative of non-linear contrast processing (Nachmias & Sansbury, 1974). Further, some studies which made use of slightly more complex stimuli than simple sinusoidal gratings have reported data that are hard to reconcile with narrowly tuned, independent spatial-frequency channels and suggest the existence of nonlinear channel interactions (e.g., Derrington & Henning, 1989; Henning, 2004; Henning, Hertz & Broadbent, 1975). Some of these issues are illustrated in Figure 1-4.

One could argue that it is difficult, if not impossible, to understand complex non-linear channel interactions without precise quantitative characterization of the contrast transduction and gain control mechanisms operating within single channels (Wichmann, 1999). Indeed, sinusoidal contrast discrimination has been studied extensively (e.g., Bird, Henning & Wichmann, 2002; Foley, 1994; Legge & Foley, 1980; Nachmias & Sansbury, 1974; Wichmann, 1999). These studies will be discussed in more detail in the following chapters. For now, it is sufficient to know that sinusoidal contrast discrimination data require a linear-nonlinear cascade model, i.e., a model of our early visual system in which a linear filtering stage is followed by a nonlinearity and internal noise. Some models place the nonlinearity in the sensory processing stage (e.g., Legge & Foley, 1980; Yang & Makous, 1995), while others leave the sensory processing linear but place the nonlinearity in the decision stage (e.g., Pelli, 1985). All these models assume that sinusoidal contrast discrimination data characterize contrast processing mechanisms operating within a single channel tuned to the signal. As will be discussed in detail in Chapter 4, recent findings suggest that even this assumption may be wrong (Henning & Wichmann, 2007).

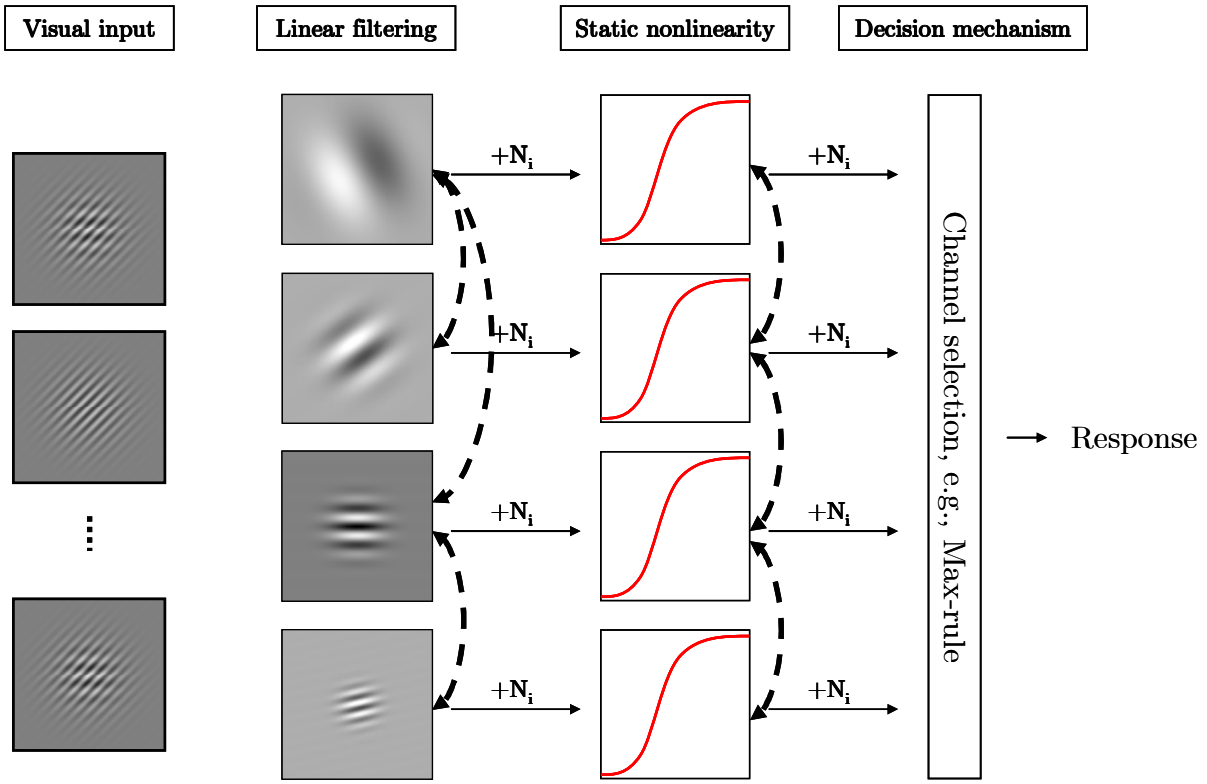

Figure 1-4. *Computational modelling of supra-threshold vision. Contrast detection and discrimination studies using stimuli that are more complex than single-component sinusoidal gratings – as illustrated under 'visual input'; in this case the observer has to detect a low-frequency horizontal grating in the presence of a high-frequency masker, having an orientation remote from the signal – have provided behavioural results that are hard to reconcile with the notion of independent and linear spatial-frequency channels (e.g., Derrington & Henning, 1989; Foley, 1994; Henning, 2004). The reader may for instance notice that the horizontal grating in the most leftward column appears to be somewhat tilted in the presence of the masker. Results like these suggest the existence of complex, nonlinear channel interactions (illustrated by the dotted double arrows).*

1.2.3 The neural substrate of spatial vision

Optical elements in the eye form an image of the visual stimulus at the level of the retina, an outgrowth of the brain also located in the eye. The retina contains several kinds of photoreceptors – which transduce the light energy into neural activity –, interneurons and ganglion cells. All together these may constitute 60 to 80 different cell types (Masland, 2001). The ganglion cells transmit visual information to the rest of the brain. Most retinal signals must pass through the lateral geniculate nucleus (LGN) to reach visual cortex. LGN has long been considered to be a passive relay

station, but this view has recently been contested (see e.g., Sherman, 2001). The first cortical visual area is the striate cortex, also called V1.

It is often argued that the neural substrate for spatial vision is likely to be found in primary visual cortex, because V1-cells share several characteristics with the psychophysical spatial-frequency channels that are not seen earlier, i.e., binocularity, directional selectivity and narrow orientation and spatial frequency selectivity (Blakemore & Campbell, 1969; DeValois & DeValois, 1988; Graham, 1989). In the tradition of Hubel and Wiesel (1959, 1962), V1-cells are often, but not always, categorized as "simple cells" or "complex cells", where simple cells essentially behave as linear spatio-temporal filters.

One of the main goals of neuroscience is to produce computational models that capture the behaviour of neurons. For the early visual system, this approach has been applied with some success for some retinal cell types (e.g., Field & Chichilniski, 2007), for some LGN cell types (e.g., Mante, Frazor, Bonin, Geisler & Carandini, 2005) and for cells in the striate cortex (e.g., Albrecht, Geisler, Frazor & Crane, 2002; Carandini & Heeger, 1994; Carandini, Heeger & Movshon, 1997; Heeger, 1992a-b; Geisler & Albrecht, 1997).

The standard view on neurons in striate cortex is well described by models that include a linear filter – which gives the cell its specific spatial and temporal selectivity – as their first stage (e.g., Carandini et al., 2005; Hubel & Wiesel, 1962; Jones & Palmer, 1987). In a subsequent stage, the filter response is squared, half-wave rectified and normalized via non-specific suppression (Carandini et al., 1997; Heeger, 1992a-b) – all these properties may be intrinsic to the feed-forward mechanisms of simple cells (Carandini, Heeger & Senn, 2002; Finn, Priebe & Ferster, 2007; Freeman, Durand, Kiper & Carandini, 2002; Miller & Troyer, 2002). This normalization model accounts for many linear and nonlinear response properties of cortical cells and is a fairly well agreed upon "standard model" for V1. It will be discussed in more detail in the following chapters. Nevertheless, much is still unknown about V1 activity – some claim up to 85 % – (Olshausen & Field, 2005).

1.2.4 Stimulus material: in defence of simplicity

The visual system has evolved to process natural images. Indeed, ultimately, the goal of visual science is to understand the computations underlying natural vision. Nevertheless, vision researchers – both physiologists and psychophysicists – have very often used and continue to use spots, bars, gratings and other mathematically elegant functions such as white noise to investigate early visual processing. These artificial stimuli differ in many aspects from natural stimuli (see Figure 1-5).

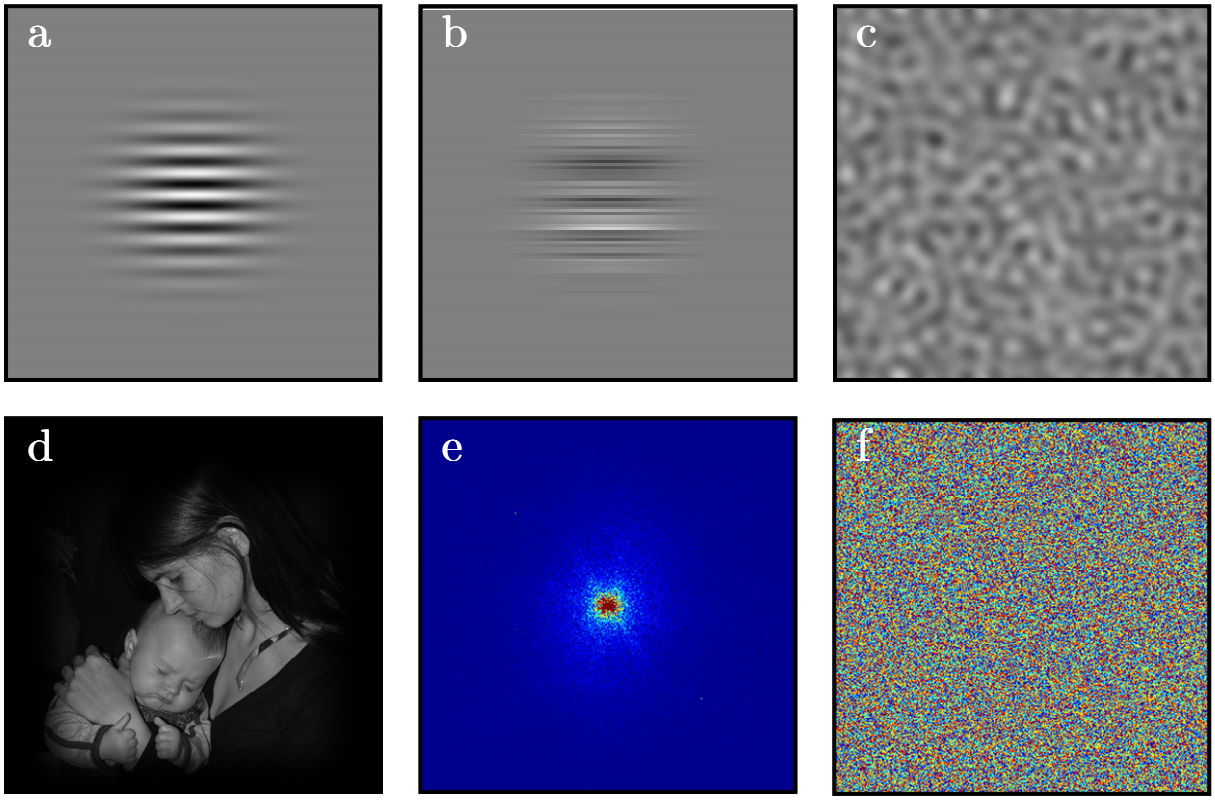

Figure 1-5. *Upper row: stimuli typically used in early vision research.* ***a.*** *A Gabor-function, having a unique spatial frequency, orientation and phase.* ***b.*** *Circularly windowed 1-D white noise, containing a multitude of randomly sampled spatial frequencies and phases.* ***c.*** *Low-pass filtered 2-D noise, containing only low spatial frequencies. Lower row: illustration of characteristics of natural images.* ***d.*** *A visual stimulus more likely to be encountered in the natural environment than any stimulus in the upper row.* ***e.*** *The Fourier power spectrum of the image shown in panel d on polar coordinates (more reddish pixels correspond to higher component amplitudes). Note that images like these typically consist of a multitude of spatial frequencies and orientations and that power decreases with spatial frequency.* ***f.*** *The phase spectrum of the image shown in panel d. Some claim that the richness of natural images is needed to develop models of visual processing, while others argue that the use of these stimuli is fraught with too much difficulty due to a multitude of correlations.*

As argued in the section on near-threshold vision, this difference is immaterial for linear visual systems. For a linear system, responses to natural images would be perfectly predictable from models developed with carefully controlled artificial stimuli. However, the early visual system is highly nonlinear, both at physiological and psychophysical level.

Recently, there has been much debate on the role natural stimuli should play in vision research (see e.g., Bex, Marechal & Dakin, 2007; Felsen & Dan, 2005; Olshausen & Field, 2005; Rust & Movshon, 2005). First, what all agree upon is that direct measurement of statistical properties of natural scenes has great potential value for advancing our understanding of visual processing (Geisler, 2008). For instance, it has been speculated that contrast-gain control mechanisms serve to produce efficient coding (Bethge, 2006; Schwartz & Simoncelli, 2001).

Opinions differ much more on the usefulness of natural stimuli in model development. To explore neurons and mechanisms with complex properties (such as those in higher cortical areas), these stimuli may be the best and perhaps only way to begin (for an example, see Vogels, 1999a-b). But should they also replace artificial stimuli for neurons and mechanisms that are better understood (such as early vision), simply because the standard models derived from the use of artificial stimuli have a limited success in predicting responses to natural stimuli? Some claim this is indeed the way to go and argue that a full understanding of the operation of complex nonlinear systems requires investigating the system in the richness of its natural environment (Felsen & Dan, 2005; Olshausen & Field, 2005; for an example in spatial vision research see Bex et al., 2007).

Others have pointed out that exploration is followed by model-building (for an example, see Op de Beeck, Wagemans & Vogels, 2001) and that models – and thus our understanding – will only improve when suspected failures of the model are investigated in a systematic and carefully controlled way (Rust & Movshon, 2005). The ultimate tool of model-building is hypothesis testing and this requires carefully controlled – and thus usually artificially constructed – stimuli.

Similarly, in spatial vision, the failure of the multi-channel model, or indeed of any model, to explain results stemming from the use of slightly more complex stimuli than sinusoidal gratings does not invalidate simple gratings as useful stimuli. To the contrary, such results indicate that the model is not good enough and needs to be improved and tested by carefully controlled stimuli. Gabor functions and white noise continue to be such stimuli (Barlow, 1956; Enroth-Cugell & Robson, 1966). Thus, while spatial vision models ultimately need to explain how we process natural images, developing and rigorously testing these models requires artificial and carefully controlled stimuli. This is the approach that will be followed in this thesis.

1.2.5 Rationale of this thesis

The central goal of this thesis is to advance our understanding of the visual processing of sinusoidal gratings embedded in visual noise. Single component gratings are perhaps the single most popular stimuli to investigate both near- and supra-threshold vision. Often, visual noise is added to these stimuli to characterize certain aspects of the visual mechanisms mediating contrast processing (e.g., Carter & Henning, 1971; Henning, 1988; Henning, Hertz & Hinton, 1981; Henning & Wichmann, 2007; Lu & Dosher, 2008; Pelli, 1985; Pelli & Farell, 1999; Solomon & Pelli, 1994). Visual noise is attractive because it is a random stimulus – and thus able to yield unbiased results –, yet its characteristics are highly controllable. (Note that the statistical properties of noise are true for the ensemble average, consisting of a large number of samples, but not necessarily for each instantiation. This stimulus variability ensures that in performing perceptual tasks, observers cannot rely on accidental local image cues provided by the noise.)

However, effects of visual noise on the processing of sinusoidal gratings are not well understood. At the physiological level, there have been very few attempts to systematically study contrast processing in the presence of visual noise. At the psychophysical level, assumptions regarding noise effects vary from excitation to inhibition (DeValois & DeValois, 1988). Consequently, conclusions inferred from experiments using noise methods may be flawed.

Here, making use of behavioural measurements and computational modelling, it will be argued that some earlier conclusions regarding contrast detection and discrimination in visual noise indeed may need to be reconsidered.

1.3 Specific background

As all communication, visual perception is intrinsically limited by noise. This noise is caused by various sources. For more than half a century, vision researchers have aimed to characterize these limitations at the overall system level by making use of external noise methods and computational models (Barlow, 1956; Nagaraja, 1964; see Lu & Dosher, 2008 for an overview and review). For detection of Gabor stimuli, the degree of complexity varies across models from a simple, two-parameter linear equation to multiple-pathway models that comprehend nonlinear transduction and several noise sources.

Despite this variety in complexity, *all* most prominent observer models[1] predict linear threshold-vs.-noise-contrast functions and parallel psychometric functions – i.e., psychometric functions are parallel on semi-logarithmic coordinates, irrespective of the external noise level – (Lu & Dosher, 2008). This is a consequence of the assumption, shared amongst these models, that the addition of external noise affects the variance of the decision statistic, but leaves the mean internal response unchanged.

In Chapter 2, this view is challenged. First, it is argued that the assumption regarding noise effects is inconsistent with the standard psychophysical model of the early visual system. This model actually predicts non-linear contrast-detection effects caused by weak external noise. Second, it is demonstrated empirically that including weak levels of noise in the external noise paradigm reveals that threshold-vs.-noise-contrast functions are indeed dipper-shaped and that psychometric functions are not

[1] As in Lu and Dosher (2008), "the most prominent observer models" refer to the *Linear Amplifier Model* (e.g., Pelli, 1981), the *Linear Amplifier Model With Uncertainty* (e.g. Pelli, 1985), the *Induced Noise Model* (e.g., Burgess & Colborne, 1988), the *Induced Noise With Uncertainty Model* (e.g., Eckstein, Ahumada & Watson, 1997) and the *Perceptual Template Model* (e.g., Lu & Dosher, 2008).

parallel on semi-logarithmic coordinates. Finally, an elaborated version of the standard contrast gain-control model is shown to fit our data reasonably well.

In Chapter 3, the hypothesis that a single mechanism underlies nonlinear effects in contrast discrimination – i.e., improved contrast detection in the presence of a low-contrast masking grating – *and* contrast detection-in-noise – i.e., improved contrast detection in the presence of subthreshold noise – is addressed. To this end, sinusoidal contrast discrimination is measured without noise and in both weak and moderate levels of noise. Making use of a full quantitative description of our data with few parameters combined with comprehensive model selection assessments, we find support for the parsimonious explanation that a single mechanism underlies both nonlinear effects.

Psychophysically measured nonlinear effects in sinusoidal contrast discrimination are typically attributed to mechanisms operating within single spatial frequency and orientation tuned channels. However, recent findings have challenged this view (Henning & Wichmann, 2007). In Chapter 4, a new model for human contrast discrimination is proposed. The encoding stage in the model is based on neurophysiological findings and models, while the decoding stage has its roots in notions from ideal observer theory. It is demonstrated that this population-code model can explain several hitherto mysterious aspects of human contrast discrimination.

In Chapter 5, some implications of the work discussed in Chapter 2-4 are addressed.

1.4 References

Albrecht, D. G., Geisler, W. S., Frazor, R. A., & Crane, A. M. (2002). Visual cortex neurons of Monkeys and Cats: Temporal dynamics of the contrast response function. *Journal of Neurophysiology, 88,* 888-913.

Banks, M. S., Geisler, W. S., & Bennett, P. J. (1987). The physical limits of grating visibility. *Vision Research, 27(11)*, 1915-1924.

Barlow, H. B. (1956). Retinal noise and absolute threshold. *Journal of the Optical Society of America, 46,* 634-639.

Bethge, M. (2006). Factorial coding of natural images: How effective are linear models in removing higher-order dependencies? *Journal of the Optical Society of America A, 23,* 1253-1268.

Bex, P. J., Marechal, I., & Dakin, S. C. (2007). Contrast gain control in natural scenes. *Journal of Vision, 7(11):12*, 1-12.

Bird, C., Henning, G. B., & Wichmann, F. A. (2002). Contrast discrimination with sinusoidal gratings of different spatial frequency. *Journal of the Optical Society of America A, 19*, 1267-1273.

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

Bruce, V., Green, P. R., & Georgeson, M. A. (1996). *Visual perception* (3rd ed.). Psychology Press.

Burgess, A. E., & Colborne, B. (1988). Visual signal detection IV. Observer inconsistency. *Journal of the Optical Society of America A, 2*(A), 617-627.

Campbell, F. W., & Green, D. (1965). Optical and retinal factors affecting visual resolution. *Journal of Physiology (London), 181*, 576-593.

Campbell, F. W., & Gubisch, R. W. (1966). Optical quality of the human eye. *Journal of Physiology (London), 186*, 558-578.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology (London), 197*, 551-556.

Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J, Dan, Y., Olshausen, B. A., Gallant, J. L., & Rust, N. C. (2005). Do we know what the early visual system does? *Journal of Neuroscience, 25(46)*, 10577-10597.

Carandini, M., & Heeger, D. J. (1994). Summation and division by neurons in primate visual cortex. *Science, 264*, 1333-1336.

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience, 17(21),* 8621–8644.

Carandini, M., Heeger, D. J., & Senn, W. (2002). A synaptic explanation of suppression in visual cortex. *Journal of Neuroscience, 22(22)*, 10053-10065.

Carter, B. E., & Henning, G. B. (1971). The detection of gratings in narrow-band visual noise. *Journal of Physiology (London), 219*, 355-365.

Derrington, A. M., & Henning, G. B. (1989). Some observations on the masking effect of two-dimensional stimuli. *Vision Research, 29(2)*, 241-246.

DeValois, R. L., & DeValois, K. K. (1988). *Spatial Vision.* Oxford University Press.

Eckstein, M. P., Ahumada, A. J., Jr., & Watson, A. B. (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, 14(A),* 2406-2419.

Enroth-Cugell, C., & Robson, J. G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology, 187*, 517-552.

Fechner, G. T. (1860). *Elemente der Psychophysik. Leipzig: Breitkopf und Härtel. (Reprinted, Bristol: Thoemmes Press, 1999).*

Felsen, G., & Dan, Y. (2005). A natural approach to studying vision. *Nature Neuroscience, 8(12)*, 1643-1646.

Field, G. D., & Chichilniski, E. J. (2007). Information processing in the primate retina: circuitry and coding. *Annual Review of Neuroscience, 30*, 1-30.

Finn, I. M., Priebe, N. J., & Ferster, D. (2007). The emergence of contrast-invariant orientation tuning in simple cells of cat visual cortex. *Neuron, 54*, 137-152.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: masking experiments require a new model. *Journal of the Optical Society of America A, 11(6)*, 1710–1719.

Frazor, R. A., & Geisler, W. S. (2006). Local luminance and contrast in natural images. *Vision Research 46,* 1585-1598.

Freeman, T. C. B., Durand, S., Kiper, D. C., & Carandini, M. (2002). Suppresion without inhibition in visual cortex. *Neuron, 35*, 759-771.

Geisler, W.S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology, 59,* 167-192.

Geisler, W. S., & Albrecht, D. G. (1997). Visual cortex neurons in monkeys and cats: Detection, discrimination, and identification. *Visual Neuroscience, 14(5)*, 897-919.

Georgeson, M. A., May, K. A., Freeman, T. C. A., & Hesse, G. S. (2007). From filters to features: Scale-space analysis of edge and blur coding in human vision. *Journal of Vision, 7(13)*, 1-21.

Graham, N. (1989). *Visual Pattern Analyzers*. Oxford University Press.

Graham, N. (1992). Breaking the visual stimulus into parts. *Current Directions in Psychological Science, 1(2)*, 55-61.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics.* New York: John Wiley & Sons, Inc.

Heeger, D. (1992a). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9*, 181-197.

Heeger, D. (1992b). Half-squaring in responses of cat striate cells. *Visual Neuroscience, 9*, 427-443.

Henning, G. B. (1988). Spatial-frequency tuning as a function of temporal frequency and stimulus motion. *Journal of the Optical Society of America A, 5*, 1362-1373.

Henning, G. B. (2004). Masking effects of low-frequency sinusoidal gratings on the detection of contrast modulation in high-frequency carriers. *Journal of the Optical Society of America A, 21(4)*, 486-490.

Henning, G. B., Bird, C. M., & Wichmann, F. A. (2002). Contrast discrimination with pulse trains in pink noise, *Journal of the Optical Society of America A, 19*, 1259-1266.

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

Henning, G. B., Hertz, B. G., & Hinton, J. L. (1981). Effects of different hypothetical detection mechanisms on the shape of spatial-frequency filters inferred from masking experiments: I. noise masks. *Journal of the Optical Society of America, 71*, 574-581.

Henning, G. B., & Wichmann, F. A. (2007). Some observations on the pedestal effect. *Journal of Vision, 7(1:3)*, 1-15.

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology (London)*, *148*,:574-591.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology (London)*, *160*, 106–154.

Jones, J. P., & Palmer, L. A. (1987). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology, 58*, 1187–1211.

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

Lu, Z.-L., & Dosher, B. A. (2008). Characterizing observers using external noise and observer models: assessing internal representations with external noise. *Psychological Review, 115(1),* 44-82.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection Theory: A user's guide*. New York: Cambridge University Press.

Mante, V., Frazor, R. A., Bonin, V., Geisler, W. S., & Carandini, M. (2005). Independence of luminance and contrast in natural scenes and in the early visual system. *Nature Neuroscience, 8(12)*, 1690-1697.

Masland, R. H. (2001). The fundamental plan of the retina. *Nature Neuroscience, 4(9)*, 877-886.

Miller, K. D., & Troyer, T. W. (2002). Neural noise can explain expansive, power-law nonlinearities in neural response functions. *Journal of Neurophysiology, 87,* 653-659.

Nachmias, J. (1972). Signal detection theory and its application to problems in vision, in *Handbook of Sensory Physiology*, vol. VII/4, pp. 56-77. Berlin: Springer Verlag.

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

Nagaraja, N. S. (1964). Effect of luminance noise on contrast thresholds, *Journal of the Optical Society of America 54(7)*, 950-955.

Olshausen, B. A., & Field, D. J. (2005). How close are we to understanding V1? *Neural Computation, 17,* 1665-1699.

Op de Beeck, H. P., Wagemans J., & Vogels, R. (2001). Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nature Neuroscience, 4(12)*, 1244-1252.

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

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America. A, 2(9)*, 1508-1532.

Pelli, D. G., & Farell, B. (1999). Why use noise? *Journal of the Optical Society of America A, 16*, 647-653.

Rust, N. C., & Movshon, J. A. (2005). In praise of artifice. *Nature Neuroscience, 8(12)*, 1647-1650.

Schade. O. H. (1956). Optical and photoelectric analogue of the eye. *Journal of the Optical Society of America, 46*, 721-739.

Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nature Neuroscience, 4(8),* 819-825.

Sherman, S. M. (2001). Tonic and burst firing: dual modes of thalamocortical relay. *Trends in Neurosciences, 24(2)*, 122-126.

Simoncelli, E. P., & Freeman, W. T. (1995). *The steerable pyramid: a flexible architecture for multi-scale derivative computation*, 2nd Annual IEEE Intl. Conference on Image Processing, Washington, DC.

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

Tanner, W. P., & Swets., J. A. (1954). A decision-making theory of visual detection. *Psychological Review, 61*, 401-409.

Vogels, R. (1999a). Categorization of complex visual images by rhesus monkeys. Part 1: behavioural study. *European Journal of Neuroscience, 11*, 1223-1238.

Vogels, R. (1999b). Categorization of complex visual images by rhesus monkeys. Part 2: single-cell study. *European Journal of Neuroscience, 11*, 1239-1255.

Wichmann, F. A. (1999). *Some aspects of modelling human spatial vision: Contrast discrimination.* Unpublished doctoral dissertation, The University of Oxford, Oxford, UK.

Wickens, T. D. (2002). *Elementary Signal Detection Theory.* Oxford: University Press.

Yang, J., & Makous, J. (1995). Modeling pedestal experiments with amplitude instead of contrast. *Vision Research, 35(14)*, 1979-1989.

Chapter 2 Some observations on contrast detection in noise

The standard psychophysical model of our early visual system consists of a linear filter stage, followed by a non-linearity and an internal noise source. If a rectification mechanism is introduced at the output of the linear filter stage, as has been suggested on some occasions, this model actually predicts that human performance in a classical contrast detection task might benefit from the addition of weak levels of noise. Here, this prediction was tested and confirmed in two contrast detection tasks. In experiment 1, observers had to discriminate a low-contrast Gabor pattern from a blank. In experiment 2, observers had to discriminate two low-contrast Gabor patterns identical on all dimensions, except for orientation (-45° vs. +45°). In both experiments, weak-to-modest levels of 2-D, white noise were added to the stimuli. Detection thresholds vary non-monotonically with noise power, i.e., some noise levels improve contrast detection performance. Both simple uncertainty reduction and an energy discrimination strategy can be excluded as possible explanations for this effect. We present a quantitative model consistent with the effects and discuss the implications.

This chapter has been published as:

Goris, R. L. T., Zaenen, P. , & Wagemans, J. (2008). Some observations on contrast detection in noise. *Journal of Vision*, *8*(9):4, 1-15

The paper can be downloaded from: http://journalofvision.org/8/9/4/

2.1 Introduction

Computational models describing the psychophysical mechanisms mediating pattern vision typically consist of a linear filter stage, followed by a static nonlinearity and a late, internal noise source (i.e., prior to the decisional stage, but after the non-linearity; Nachmias, 1989). Further, detection threshold studies, i.e., experiments that involve barely visible, low-contrast stimuli, have provided much evidence consistent with a model of our visual system in which visual processing is accomplished in many, relatively independent spatially localised spatial-frequency and orientation-selective filters (Blakemore & Campbell, 1969; Campbell, Carpenter & Levinson, 1969; Campbell & Robson, 1968; DeValois & DeValois, 1988; Graham & Nachmias, 1971; Henning, Hertz & Hinton, 1981; Stromeyer & Julesz, 1972; Wilson, McFarlane & Phillips, 1983). The psychophysical receptive field profiles of these filters, while slightly asymmetric on double logarithmic coordinates (Henning, 1988; Henning et al., 1981), are nonetheless typically approximated by Gabor functions (Figure 2-1). Though not necessarily an 'optimal' stimulus (Watson, Barlow & Robson, 1983; but see Henning, Derrington & Madden, 1983), Gabors are popular stimuli to explore contrast detection because they are band-limited in spatial-frequency and localised in space (Daugman, 1985).

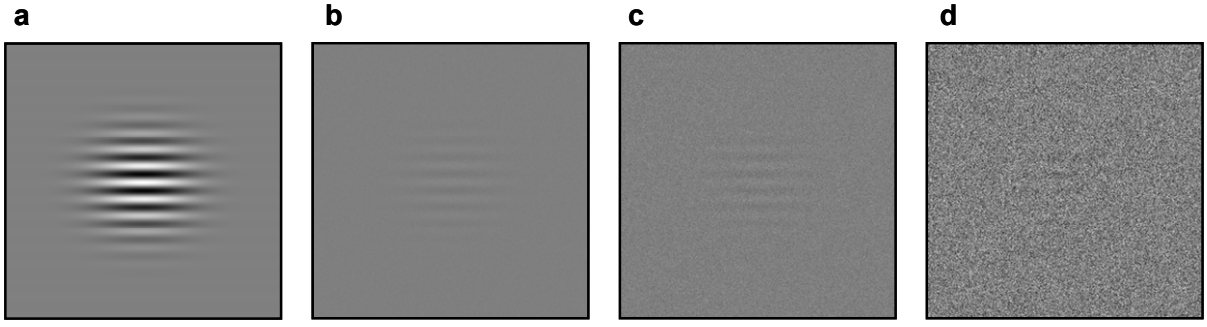

Figure 2-1. *These stimuli ought to be viewed from a distance of 60 cm.* ***a.*** *A Gabor function at maximal contrast. This stimulus had a spatial frequency of 7 c/deg and resembles the stimulus used in experiment 1.* ***b.*** *The same stimulus at 5% signal contrast. When presented for about 25 ms, this would be around detection threshold for the majority of observers. Weak, 2-D white noise has been added to this display. When viewed from a distance of 60 cm, the noise power matches the lowest level of noise used in our experiments.* ***c.*** *The same stimulus at 5% signal contrast, embedded in the 'optimal' level of noise used in our experiments.* ***d.*** *The same stimulus at 5% signal contrast, embedded in the highest level of noise used in our experiments.*

To gain insight in the visual system's processing at suprathreshold contrasts, sinusoidal contrast discrimination has been studied extensively (e.g., Bird, Henning & Wichmann, 2002; Foley, 1994; Foley & Chen, 1997; Foley & Legge, 1981; Gorea & Sagi, 2001; Henning & Wichmann, 2007; Kontsevich, Chen & Tyler, 2002; Legge, 1981; Legge & Foley, 1980; Legge, Kersten & Burgess, 1987; Nachmias & Sansbury, 1974; Wichmann, 1999; Yang & Makous, 1995). The main finding of these studies is the pedestal or dipper effect: A sinusoidal grating is much more detectable when added to a low-contrast masking or pedestal grating of identical spatial frequency, orientation, phase, and temporal duration. Once the contrast of the pedestal grating exceeds a certain value, discrimination thresholds rise in a Weber's law-like fashion.

To explain this dipper-shaped threshold-vs.-contrast function, contrast perception models typically include a nonlinear post-filter stage (e.g., Legge & Foley, 1980; Foley, 1994; Foley & Chen, 1997; Wichmann, 1999; Yang & Makous, 1995). The response expansion that is believed to occur at low contrasts (and thus weak filter responses) is sometimes thought to underlie the initial threshold decrease although this view has recently been questioned (Henning & Wichmann, 2007).

Because human performance is not perfect (i.e., both in contrast detection and discrimination experiments, human thresholds differ from ideal observer thresholds), psychophysical models must either assume or postulate limited (neural) efficiency and/or an internal noise source. There has been some debate regarding the question whether the crucial, performance limiting noise source is to be found in the early or rather in the later processing stages (e.g., Henning, Bird & Wichmann, 2002; Nachmias, 1989; Pelli, 1991). Furthermore, this might differ for contrast processing in detection and discrimination circumstances (Henning et al., 2002). It is thus not surprising that models describing the psychophysical mechanisms underlying detection in noise have considered several possible locations for the main, internal noise source (Lu & Dosher, 2008). Nevertheless, the standard model of the psychophysical mechanisms mediating pattern vision that proposes a linear filter stage, a static nonlinearity and a late internal noise source, is fairly consistent with most contrast discrimination data.

There is a clear analogue for this model at the neural level. The neural substrate for contrast processing is likely to be found in primary visual cortex (e.g., DeValois & DeValois, 1988). The classical view on neurons in these early stages of visual processing is well described by models that include a linear filter as their first stage (e.g., Carandini et al., 2005). In a subsequent stage, contrast normalization via non-specific suppression takes place, leading to contrast-response functions that can be described by Naka-Rushton functions (e.g., Albrecht, Geisler & Crane, 2003; Carandini, Heeger & Senn, 2002; Freeman, Durand, Kiper & Carandini, 2002; Geisler & Albrecht, 1995, 1997; Heeger, 1992a). It has been speculated that this type of image analysis serves, amongst other goals, to produce efficient coding (Bethge, 2006; Schwartz & Simoncelli, 2001).

One crucial element of these models has not been mentioned yet. Because cells cannot have negative firing rates and linear filters do give negative responses, a mechanism removing negative responses must be introduced before the normalization stage (e.g., Carandini et al., 2005; Heeger, 1992b). Different suggestions have been made, amongst which half-, full- and over-rectification, and half- and full-squaring (in the latter case, (rectified) on- and off-cells may be sources of both halves of the response). All achieve the removal of negative responses. Similar rectification mechanisms have successfully been introduced in psychophysical contrast perception models (e.g., Foley, 1994; Legge & Foley, 1980). One consequence of the rectification of a linear filter's output is that it leads to a monotonically rising filter response as a function of noise power. In other words, if a rectification mechanism is implemented, externally added noise not only affects internal response variability, but also raises the mean response to a stimulus, at least at low contrast levels. Because weak filter responses are followed by response expansion in pattern vision models, these models actually predict that contrast detection performance should improve slightly if weak levels of white noise are added to the stimuli. Due to the introduced stimulus variability and the random phase relation between signal and noise, this improvement is likely to be relatively mild compared to the classical pedestal effect. At higher noise levels, thresholds are expected to rise linearly (see Figure 2-2).

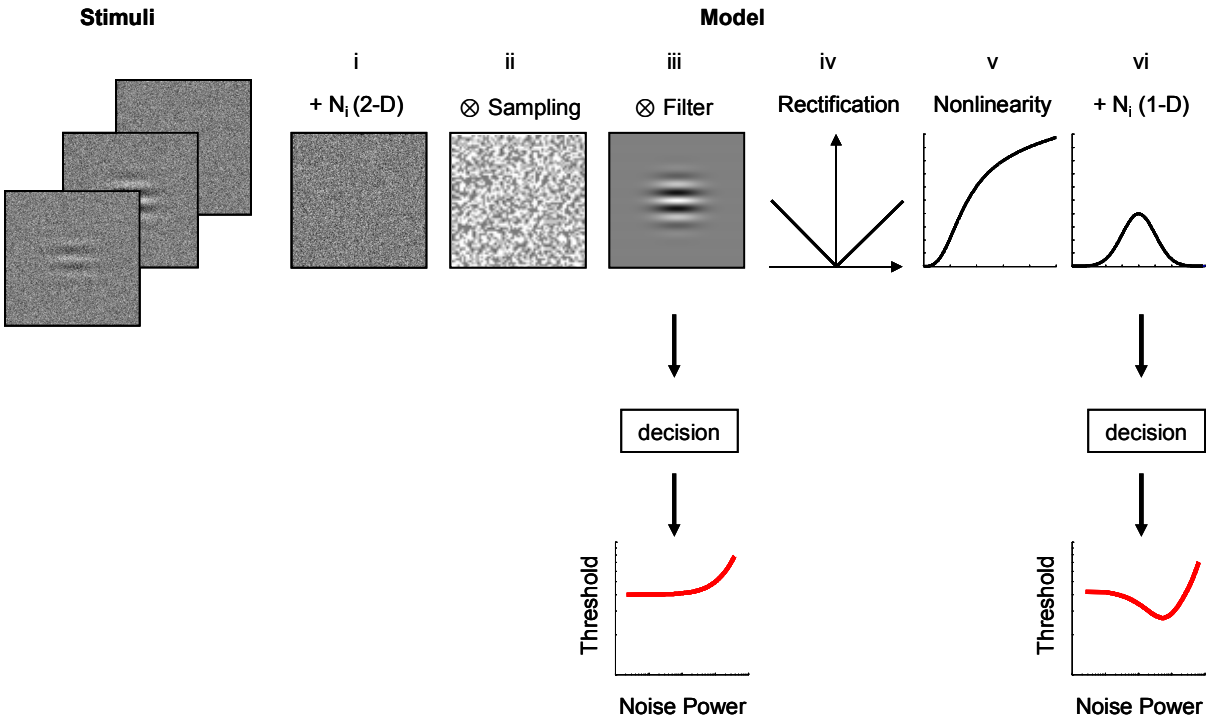

Figure 2-2. *Illustration of the stimuli (left) and model components (right) discussed in the text. Stimuli with a narrow-band spatial-frequency spectrum are often assumed to be processed in a single spatially localised, spatial-frequency and orientation-selective filter. Ideally, an observer uses a filter that is an exact template of the signal (iii). Human deviations from ideal observer performance are typically interpreted as stemming from internal noise sources (i and vi) on the one hand and limited efficiency or sampling (ii) on the other hand (this refers to using only a sample of the available image information, for instance only the image pixels corresponding to the white pixels in ii). Early noise has the same dimensionality as the signal (i), while late noise is introduced prior to the decision, where the signal representation is 1-D (vi). If the signal representation at the decisional stage is directly based on the (rectified) filter response, a linear threshold-vs.-noise-power function is to be expected. To explain the pedestal effect seen in contrast discrimination, models of pattern vision must include a nonlinear post-filter stage, for instance a Naka-Rushton transformation (v). If the output of the human psychophysical filters mediating detection of a Gabor stimulus in noise is subject to rectification (iv), nonlinear transduction (v) and dominant late noise (vi) prior to the decisional stage, a dipper-shaped threshold-vs.-noise-power function is to be expected. The rectifying and Naka-Rushton nonlinearities are presented as different stages in this figure, to disentangle their effects on the internal signal representation. They could, however, be collapsed into a single positively accelerated nonlinearity. Further, as mentioned in the text, the rectification need not be full-wave.*

The combination of rectification and response expansion has typically not been included in contrast perception models describing detection in noise, so that a linear

threshold-rise is predicted at all noise levels (e.g., Dosher & Lu, 1998; Lu & Dosher, 2008; Pelli, 1985; Pelli & Farell, 1999), as illustrated in Figure 2-2. Exploration of detection in weak noise can thus be seen as a way to test whether both rectification and response expansion are necessary components of the standard early vision model of detection in noise.

Detection in noise has been studied on many occasions, and all but one study (Blackwell, 1998) reported monotonically rising detection thresholds as a function of noise power (e.g., Carter & Henning, 1971; Dosher & Lu, 1998; Pelli, 1985; Pelli & Farell, 1999). This inconsistency might simply be a consequence of the different noise levels considered in these studies: Blackwell (1998) used lower noise levels than the other studies mentioned above. She reported small facilitation effects in a classical detection-in-noise task, i.e., the discrimination of a low-contrast signal from a uniform field of mean luminance. The performance improvement was interpreted as resulting from uncertainty reduction, i.e., certain noise levels reduce temporal uncertainty about the stimulus.

In the experiments reported here, we measured detection thresholds in the presence of weak levels of 2-D, white noise while temporal and spatial uncertainty regarding stimulus presentation was minimized by means of a high contrast visual marker and a constant temporal task profile. Feedback was used to induce good performance. Stimulus presentation time was very short to obtain relatively high detection thresholds (around 5% Michelson contrast) and avoid artefact effects due to the limited luminance sampling of our CRT. Finally, to rule out the possibility that the performance improvement found in a classical detection task (i.e., experiment 1) was due to a strategy based on global energy discrimination instead of truly enhanced signal visibility, we also measured contrast thresholds in an additional experiment. In this experiment, subjects had to discriminate two Gabors identical on all dimensions except that their orientation differed by 90°. It has been reported that orientation discrimination results in similar contrast threshold estimates as classical signal detection (e.g., Solomon & Pelli, 1994). In both experiments a decrease of contrast thresholds at certain weak noise levels was observed.

2.2 Methods

2.2.1 Equipment

The experiments were run using a Power Macintosh G4 computer, with the software packages MATLAB (Mathworks, Natick, MA) and PSYCHTOOLBOX (Brainard, 1997; Pelli, 1997). Luminance was measured with a KONICA MINOLTA CS-100 Spot Chroma Meter. Gamma correction used an 8-bit lookup table and ensured that the monitor was linear over the entire luminance range used in the experiments. The stimuli were presented on a SONY Trinitron GDM-FW900 monitor, with a spatial resolution of 1920 × 1440 pixels and a temporal resolution of 75 Hz. The experiment was run in a darkened room and the screen's mean grey background luminance was set to 42 cd/m^2. Viewing distance was 120 cm, leading to a pixel-size of .009° of visual angle.

2.2.2 Observers

Three observers participated in experiment 1 (R.V., B.B., and R.G.), and three observers participated in experiment 2 (I.P., H.H., and E.G.). All were well practiced with the task and stimuli before data collection began and had normal or optically corrected-to-normal vision. All observers, except for R.G., were naïeve to the purpose of the experiment.

2.2.3 Stimuli

2.2.3.1 Signal stimuli

The Gabor stimuli consisted of orientated sinusoidal gratings with a spatial extent of 2.35° of visual angle and a spatial frequency of 7 c/deg, which were then multiplied by a two-dimensional spatial Gaussian envelope with sigma equal to 0.27°. In experiment 1, the Gabor stimulus was the target or signal, orientated horizontally (0°). The non-target stimulus was a grey field of the same mean luminance as the signal. In experiment 2, the target stimulus was arbitrarily defined as a Gabor pattern orientated +45° (to the right); an otherwise identical Gabor pattern orientated –45° (to the left) was defined as the non-target stimulus. These angles were chosen to avoid effects of pixel contamination (due to pixel contamination, the

effective contrast of vertically orientated gratings may be lower than the contrast of horizontally orientated gratings). Further, contrast sensitivity does not differ for these oblique orientations. Examples of the Gabor pattern used in Experiment 1 are shown in Figure 2-1.

2.2.3.2 Background noise stimuli

For each stimulus presentation a fresh noise sample, of which each pixel luminance value was sampled from a Gaussian distribution centred at mean luminance, was generated. Ten levels of white noise were used, ranging between 0 and 38×10^{-7} deg^2 noise power spectral density. Noise power spectral density is defined as the variance in luminance (relative to the space-average luminance) multiplied with the pixel area, expressed in visual degrees squared. It represents the average power at the different frequencies present in the noise. The maximal amount of clipping (i.e., pixels set to the minimal or maximal luminance values because of the limited dynamical range of the monitor and video card) at the highest noise level was below 2%. Estimates of the effective images ensured that the nonlinear monitor operations (i.e., power saving functions, gamma correction, luminance rounding and the gamma function) did not distort the spectral properties of the noise stimuli.

2.2.3.3 Procedure

In both experiments 1 and 2, a temporal two-alternative-forced-choice (2AFC) task was used. Stimulus presentation time was approximately 25 ms (presentation time was doubled for subject E.G., and halved for subjects B.B. and R.G. to allow them to reach similar thresholds as the other subjects). Each stimulus presentation, the stimulus was surrounded by a red square (2.35°) at maximal luminance. This square appeared and disappeared with the stimulus onset and offset. All trials started with the 266 ms presentation of a grey field. The first stimulus presentation was followed by an inter-stimulus-interval (ISI) of 466 ms, and then followed by the second stimulus presentation. The ISI was thus about 20 times longer than the stimulus duration and more than 3 times as long as the temporal impulse response (Graham, 1989). Response time was limited to 1,000 ms, and was indicated by a grey field of mean luminance, presented after the second stimulus presentation. After the response

screen, a screen with a green square (2.35°) appeared for about 750 ms. During this presentation, auditory feedback indicated which interval contained the signal (low tone of 1,000 Hz: first interval; high tone of 2,000 Hz: second interval). Trials in which responses fell outside the one second answer interval (about 2% of the trials) were repeated. On these occasions, the interval containing the signal was selected randomly again, and the noise was also refreshed again. After a few training blocks, participants were very familiar with this steady temporal task profile.

To obtain psychometric functions, seven Michelson contrast levels were tested at each noise level. The Michelson contrast of an image is defined as the maximal luminance minus the minimal luminance divided by their sum, and thus ranges between 0 (no spatial luminance variation) and 1 (black and white present in the image). In experiment 1 an interval contained either a uniform field of mean luminance or a non-zero contrast Gabor stimulus. In experiment 2, all stimuli had the same contrast within a trial but, as mentioned above, the stimuli were orientated either +45° or –45°.

Within each block of 50 trials, the contrast and noise levels were randomised. During an entire block a red square of 3.5° of visual angle, surrounding the stimuli, at half of the maximal contrast was presented all the time. Subject I.P., E.G., R.V., B.B. and R.G. completed 7,000 trials in total; subject H.H. completed 4,000 trials in total.

2.3 Results

The results are summarised in Figure 2-3. 75% correct Michelson contrast detection thresholds are plotted as a function of noise power spectral density. The upper row shows the results for two observers (HH and RV), the lower row shows the results averaged across observers in experiment 1 (left) and 2 (right). Thresholds vary non-monotonically with the external noise level: Before increasing, thresholds reach a minimum at a particular noise level. Stated differently, some noise levels improve human contrast detection. This effect is present for all observers, in both experiments.

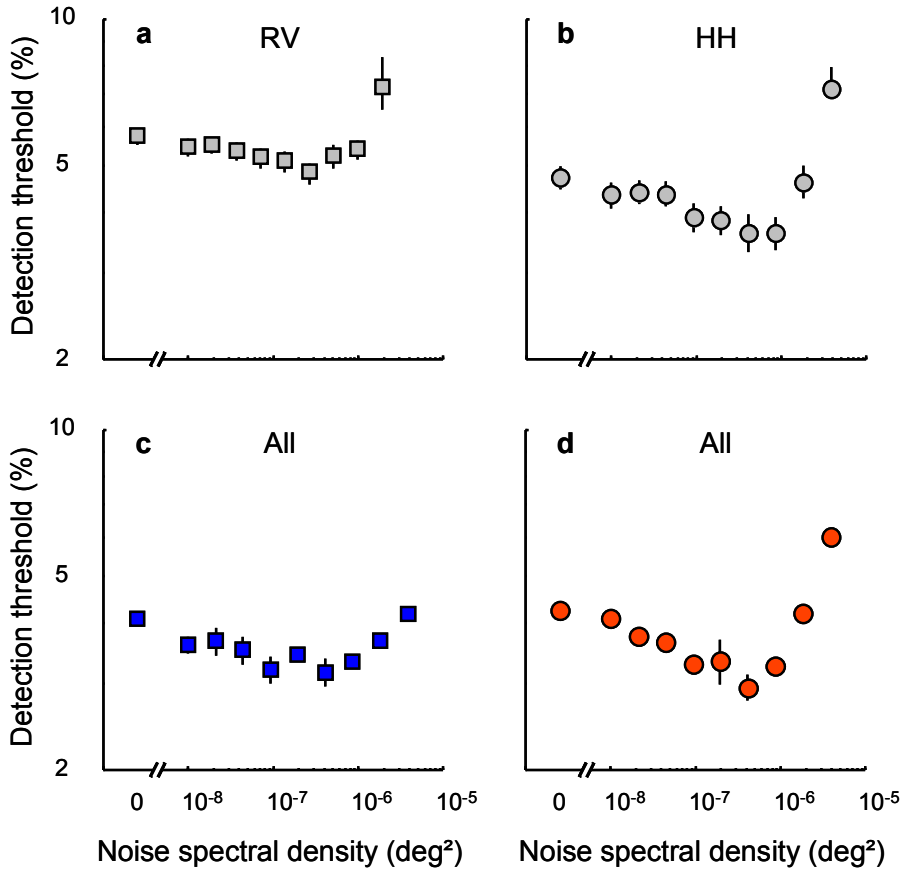

Figure 2-3. *75% correct contrast detection thresholds as a function of noise spectral density on double logarithmic coordinates. The most leftward point denotes the noiseless detection threshold. Error bars, where visible, show estimates of ± 1 SD, calculated across observers for panel* ***c*** *and* ***d****. The left column shows results for experiment 1, the right column results for experiment 2.* ***a.*** *Results for observer RV.* ***b.*** *Results for observer HH.* ***c.*** *Results for the signal detection task, averaged across two observers (RG and BB; the data of RV have not been used for panel* ***c.*** *because the noise levels used for this observer were slightly different).* ***d.*** *Results for the orientation discrimination task, averaged across three observers (IP, EG and HH).*

As can be seen in Figure 2-4, the depth of the dip at 75% correct varies somewhat between observers, similar to the classical pedestal effect. To estimate the average maximal threshold reduction as a function of noise, thresholds for each observer were first normalised by their noiseless detection threshold. This way, the noiseless detection threshold equals one and all other thresholds are expressed as products of this detection threshold. These normalised thresholds were then averaged over observers. Via the same bootstrap procedure used to estimate confidence intervals for each participant, the average maximal 75% correct threshold reduction was estimated

to be of a factor 1.35, with the 95% confidence interval ranging from 1.25 to 1.46. At 60% correct, the threshold reduction factor was estimated to be 1.63 (95% C.I.: 1.40 - 1.89) and at 90% correct, this factor equaled 1.14 (95% C.I.: 1.06 - 1.23).

The strength of the noise benefit thus depends on the performance level considered, as is the case for the classical pedestal effect seen in contrast discrimination. However, these values are considerably smaller than the improvement factors usually observed in classical pedestal experiments. Estimates from a contrast discrimination study, making use of the same signal and procedure as reported here, suggest an average improvement factor of 3.44 at 75% correct (Goris, Wagemans & Wichmann, 2008), which is in line with other estimates (e.g., Bird, Henning & Wichmann, 2002; Henning & Wichmann, 2007). Nevertheless, the same mechanism that underlies the dipper effect in contrast discrimination might give rise to the noise benefit observed here. The pedestal effect depends on the phase relation between signal and pedestal – the dipper effect disappears with signal and pedestal 90° out-of-phase; a phase difference of 180° leads to a 'bumper effect' – (e.g., Yang & Makous, 1995). Therefore, with random phase relation between signal and noise, only a slight benefit is expected.

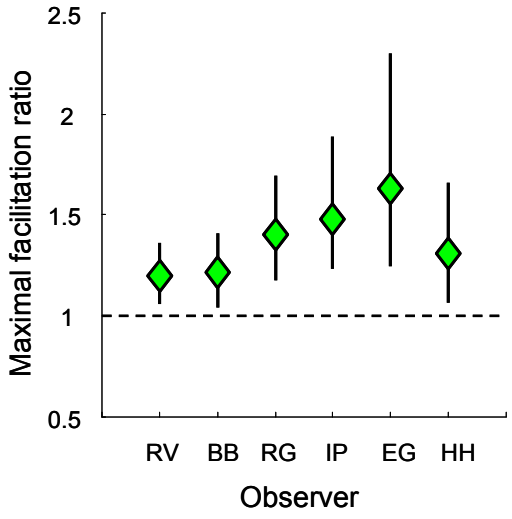

Figure 2-4. *Maximal facilitation ratio at the 75% correct detection threshold for each observer. The green symbols show the average maximal facilitation ratio, error bars indicate the 95% confidence interval. To estimate these quantities, 100,000 threshold estimates for each condition were generated by means of a bootstrap-procedure, as described in Wichmann and Hill (2001b).*

The noise facilitation effect appears to be somewhat stronger in experiment 2 than in experiment 1. This may be due to interobserver variability, which is relatively high (see Figure 2-4). On the other hand, this difference may also be due to modified strategies or stimulus representations, caused by the changed stimulus conditions in experiment 2 (e.g., the overall energy between blank stimuli and orientated Gabor stimuli differs).

Given that our experiments were designed to minimize uncertainty about space and time of the stimulus presentation by means of visual markers and a steady temporal task profile, a straightforward uncertainty reduction effect of some noise levels seems not to be a likely explanation for these results – uncertainty models of pattern vision predict no noise benefit in contrast detection (Lu & Dosher, 2008; Pelli, 1985). Furthermore, the similarity between the data of experiment 1 and 2 excludes the possibility that a global energy-discrimination strategy underlies these results. Though this explanation is not inconsistent with the results of experiment 1, the results of experiment 2 rather suggest that stimulus identity was retained.

Altogether, these results seem to be consistent with the standard psychophysical model of pattern vision, wherein a linear filter stage is followed by a rectification mechanism, expansion of weak filter responses due to a static nonlinearity and a late internal noise source. We provide quantitative examination of this suggestion in the next section.

2.4 Model: Equations, Fitting, Evaluation

2.4.1 Equations

We elaborate the standard divisive gain-control model as formalised by Wichmann (1999) for sinusoidal contrast discrimination to describe sinusoidal contrast detection in noise. In order to do so, the gain-control model is extended with two parameters typically used by linear detection in noise models, i.e., early noise and sampling.

Further, the filter stage – which need not be specified explicitly for sinusoidal contrast discrimination – is chosen to consist of (optimal) template matching (e.g., Lu & Dosher, 2008). Observer's responses are modelled within the framework of signal detection theory (SDT, Green & Swets, 1966). Three hypothetical stages are specified in SDT-models: First, a *stimulus theory* describes how a transduction mechanism maps physical stimuli to internal states, second, a probabilistic *theory of internal states* describes the probability distribution of the internal states that results from repeated presentation of the same stimulus, and finally a deterministic *response theory* describes a decision rule that maps internal states to a response.

2.4.1.1 Stimulus theory

As in some linear detection in noise models, the linear part of the transduction mechanism consists of early signal-independent or additive 2-D noise, σ^2_{add}, image sampling or calculation efficiency, k, and template matching (Lu & Dosher, 2008). The parameter k expresses the proportion of available information used by the observer and ranges between 0 and 1. Cross-correlating the noisy, sampled input image, $I_{sampled}$, with an optimal signal template, T_{signal}, transforms the 2-D input stimuli to 1-D responses, R_k, as given by equation 2-1.

$$R_k = \iint T_{signal}(x,y) I_{sampled}(x,y) dxdy \,. \qquad (2\text{-}1)$$

Subsequently, this filter response is rectified. The effects of sampling and the rectification on the mean internal representation, $|R_k|$, were estimated via simulations with the noise and signal contrast levels used in our experiments as input. The scale of these responses depends on the image size used and therefore these responses were normalised by the filter response to a full contrast signal so that filter responses to a noiseless, unsampled signal became identical to the Michelson contrast of that signal.

As explained above and illustrated in Figure 2-2, the aforementioned model components give rise to a linear relationship between image contrast and internal contrast representation. To describe the nonlinear mapping of stimulus contrast to

internal contrast representation, R(C), the second part of the transduction mechanism consisted of the three parameter Naka-Rushton function (free parameters α, β, and p), which is illustrated in Figure 2-2v and given by equation 2-2.

$$R(C) = \alpha \frac{C^p}{\beta^p + C^p}. \tag{2-2}$$

The transduction mechanism is thus fully determined by specifying the sampling (k) and the parameters of the Naka-Rushton equation (α, β and p). Equation 2-3 expresses this transduction, t, as a function of the signal contrast (C) and the effective total noise spectral density (N_{total}) given a certain sampling value k. Because the early, internal noise is additive, N_{total} is the sum of the external noise level σ^2_{ext} and the early noise σ^2_{add}.

$$t(C, N_{total}|k) = \alpha \frac{|R_k(C, N_{total})|^p}{\beta^p + |R_k(C, N_{total})|^p}. \tag{2-3}$$

It is important to note that the rectified filter responses, $|R_k|$, used in the expansive, i.e., the nominator, and the compressive, i.e., the denominator, parts of the Naka-Rushton function were the same. Although some evidence points to the existence of a broadly tuned contrast gain control pool (e.g., Foley, 1994; Holmes & Meese, 2004), we opted to use only within-channel suppression in this model to avoid an increase of the number of free parameters.

2.4.1.2 Theory of internal states

The variability of internal states is determined by internal noise on the one hand and the use of stochastic stimuli on the other hand. Three free parameters were used to describe the late 1-D internal noise, assumed to be Gaussian and having both a signal-independent (free parameter σ^2_{late}) and signal-dependent source (free parameters γ and ξ, with ξ a proportional constant and γ an exponent).

The external noise used in our experiments will inevitably contribute to the total variability of the internal representation. Because the Naka-Rushton function is – in general – nonlinear, the effect of its parameter values on the variability of the internal representation has to be estimated by means of simulations. To obtain estimates of the variance of the noise representations, we ran simulations with the noise levels used in our experiments as input. Equal variance of the signal and noise representations was assumed as a first approximation. We varied k, p and β, and fitted descriptive functions to the simulated variances. These descriptive functions allowed us to formalise the full model behaviour.

2.4.1.3 Response theory

As is standard, it was assumed that the observer's response ('interval 1' or 'interval 2') is determined by the stimulus interval that led to the highest internal state. The model equations (see equation 2-4–2-6) were arranged to express percent correct, $p(C, N_{total})$, as a function of the signal contrast (C) and the effective total noise spectral density (N_{total}) in a 2AFC-task.

$$p(C, N_{total}) = \int_0^{\infty} \frac{1}{\sqrt{2\pi g(C, N_{total})}} e^{-\frac{(z - f(C, N_{total}))^2}{2g(C, N_{total})}} dz \,, \tag{2-4}$$

where z is a dummy variable, and $f(C, N_{total})$ and $g(C, N_{total})$ are given by

$$f(C, N_{total}) = \alpha \left(\frac{\overline{\left(\left|R_k\left(N_{total}C\right)\right|\right)^p}}{\beta^p + \overline{\left(\left|R_k\left(N_{total}C\right)\right|\right)^p}} - \frac{\overline{\left(\left|R_k\left(N_{total}\right)\right|\right)^p}}{\beta^p + \overline{\left(\left|R_k\left(N_{total}\right)\right|\right)^p}} \right), \tag{2-5}$$

and

$$g(C, N_{total}) = 2\sigma_{late}^2 + 2\left(VAR\left(\alpha \frac{\left(\left|R_k(N_{total})\right|\right)^p}{\beta^p + \left(\left|R_k(N_{total})\right|\right)^p} \right) + \zeta^2 \left(\left(\frac{\alpha\left(\left|R_k(N_{total}C)\right|\right)^p}{\beta^p + \left(\left|R_k(N_{total}C)\right|\right)^p} \right)^{2\gamma} + \left(\frac{\alpha\left(\left|R_k(N_{total})\right|\right)^p}{\beta^p + \left(\left|R_k(N_{total})\right|\right)^p} \right)^{2\gamma} \right) \right). \tag{2-6}$$

In sum, by combining simulations and analytical descriptions we obtained parametric descriptions of performance in a 2-AFC task as a function of eight free parameters (σ^2_{add}, k, α, β, p, σ^2_{late}, γ and ξ) and external noise level N and signal contrast C.

2.4.2 Fitting

One of the eight parameters can be arbitrarily set to any value. To follow the usual convention, σ^2_{late} was taken to be 1, resulting in a seven free parameter model. An additional free parameter λ ("lapse rate") was introduced in the fitting of the model to avoid biased parameter estimates (Wichmann, 1999; Wichmann & Hill, 2001a). Priors were introduced for each parameter to constrain estimates to realistic values.

To find the surface $p(C, N_{total})$ that maximizes the likelihood that the data were generated from a process with success probability given by $p(C, N_{total})$, the log-likelihood of the surface $p(C, N_{total})$ given the parameters (σ^2_{add}, k, α, β, p, γ, ξ and λ) was maximized using purpose-written software in MATLAB (*fminsearch,* which makes use of the Nelder-Mead simplex search method). The log-likelihood of the surface $p(C, N_{total})$ given parameter vector θ, containing $\{\sigma^2_{add}, k, \alpha, \beta, p, \gamma, \xi, \sigma^2_{late}$ and $\lambda\}$ with $\sigma^2_{late} = 1$ is given by equation 2-7:

$$l(\theta) = \sum_{j=1}^{Z} \sum_{i=1}^{K_j} \log \binom{n_{ji}}{y_{ji} n_{ji}} + y_{ji} n_{ji} \log\left(p\left(C_{ji}, N_j; \theta\right)\right) + \left(1 - y_{ji}\right) n_{ji} \log\left(1 - p\left(C_{ji}, N_j; \theta\right)\right), \qquad (2\text{-}7)$$

with n_{ji} the number of trials (block size) measured at noise level N_j and signal contrast C_{ji} and y_{ji} the proportion of correct responses in that condition. Because the problem is non-convex due to λ, a multistart procedure with pseudo-randomly selected initial values was used to find the probable global minimum for each participant.

2.4.3 Evaluation

To evaluate model fits we considered the overall distance between model prediction and data and the presence of systematic errors in the residuals. Quality of the overall fit can be assessed by judging total deviance (see equation 2-8), i.e., the log-likelihood ratio of the saturated model and the best fitting model (the saturated model is the model with no residual error between model predictions and data). What deviance does not assess, however, are systematic trends in the deviance residuals (see equation 2-9), i.e., the agreement between individual data points and the corresponding model prediction. For binomial data, deviance is expressed by equation 2-8.

$$D = 2\sum_{j=1}^{Z}\sum_{i=1}^{K_j}\left\{ n_{ji} y_{ji} \log\left(\frac{y_{ji}}{p\left(C_{ji}, N_j\right)} \right) + n_{ji}(1-y_{ji}) \log\left(\frac{1-y_{ji}}{1-p\left(C_{ji}, N_j\right)} \right) \right\}, \qquad (2\text{-}8)$$

This statistic indicates how well a model describes data. Asymptotically, it can be shown to be χ^2-distributed, with degrees-of-freedom equal to the number of data blocks minus the number of free parameters if the model is correct and the observer's behaviour were perfectly stationary during the whole experiment (such an observer would thus generate truly binomially distributed data). Often, due to a variety of reasons, this is not the case. Responses of non-stationary observers are more variable than binomially distributed data and thus lead to higher deviances (overdispersion).

Wichmann (1999) has shown that, due to the typically relatively small number of measurements, the asymptotically derived distributions often fail to approximate the real distribution of D for psychophysical data-sets. The real distribution of D can be estimated easily by means of Monte Carlo simulations. As suggested by Wichmann (1999), we estimated the distribution of D for each model fit by means of 10,000 simulated data-sets for an observer whose correct responses in our experiment are binomially distributed as specified by the model fit. From these simulations, we derived critical values for each reported fit.

Each deviance residual d_i, is defined as the square root of the deviance value calculated for data point i in isolation, signed according to the direction of the arithmetic residual $y_i - p(C_{ji}, N_i)$. For binomial data, this is expressed by equation 2-9.

$$d_{ji} = \mathrm{sgn}\left(y_{ji} - p\left(C_{ji}, N_j\right)\right)\sqrt{2\left[n_{ji}y_{ji}\log\left(\frac{y_{ji}}{p\left(C_{ji}, N_j\right)}\right) + n_{ji}\left(1-y_{ji}\right)\log\left(\frac{1-y_{ji}}{1-p\left(C_{ji}, N_j\right)}\right)\right]}. \quad (2\text{-}9)$$

Note that $D = \sum_{j=1}^{Z}\sum_{i=1}^{K_j} d_{ji}^2$, as for RMSE. Systematic trends in deviance residuals indicate a systematic misfit of the model.

2.5 Model: Results

For each observer, the eight free parameter gain-control model was fitted to all data. A detailed example of such a fit to the data of observer IP and RV is shown in Figure 2-5. Both the psychometric functions derived from the model fit (full red lines) and Weibull functions fitted to the data (dotted red lines) are plotted. The gain-control model has roughly 1 free parameter per noise level (i.e. eight free parameters for ten noise levels) as compared to the four free parameters per noise level used by the Weibull function (Wichmann & Hill, 2001a). Nevertheless, the fits to the data are relatively similar. The more parsimonious gain-control model thus seems to give a reasonable approximation of these data.

The whole model fit is shown in Figure 2-6 for all observers. Performance, indicated by colour, is plotted as a function of noise spectral density and signal contrast. The thresholds derived from the model fits are dipper-shaped as a function of noise spectral density. Furthermore, the threshold reduction depends on performance level considered – being stronger at lower performance levels –, thus capturing the same trend seen in the raw data. Table 2-1 lists the parameter estimates and normalised total deviance for all observers.

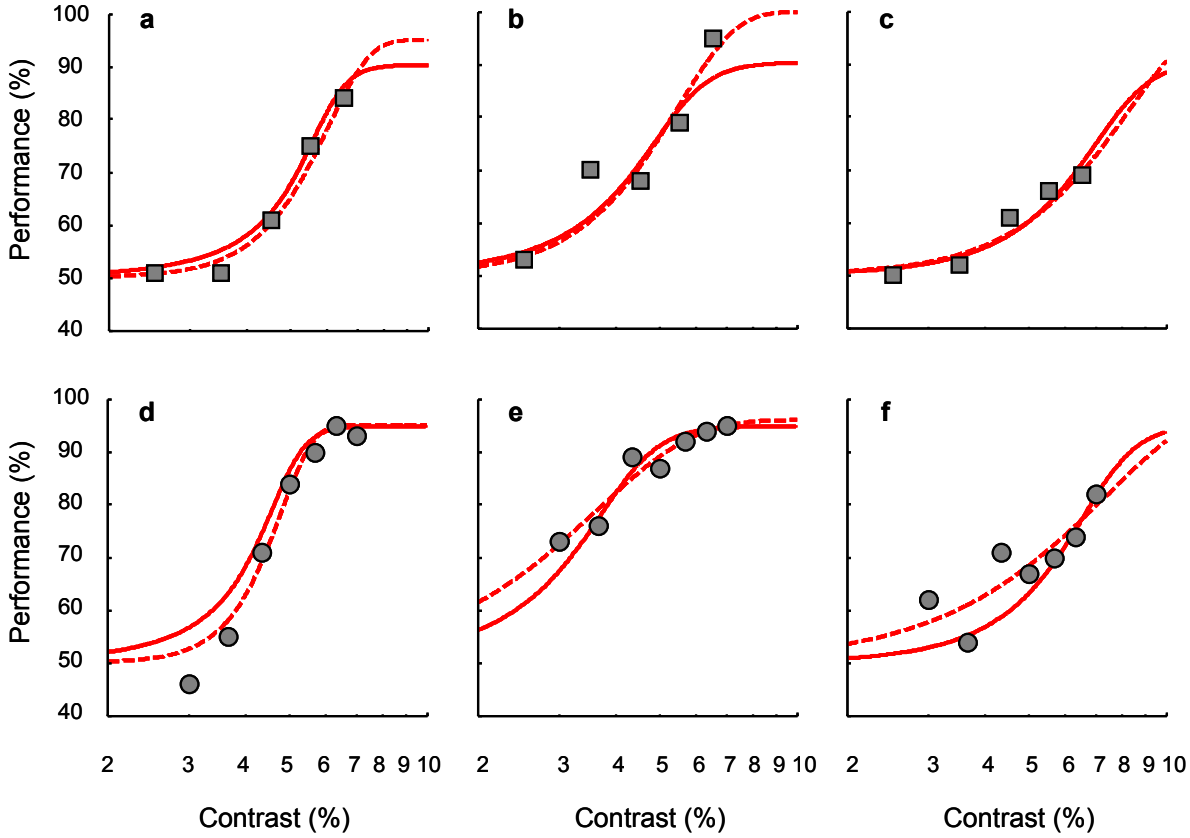

Figure 2-5. *Performance as a function of signal contrast for observers RV (squares, upper row) and IP (circles, lower row) at three different noise levels: no noise (left), 'optimal' noise (middle) and the highest noise level used in the experiment (right). Grey circels indicate measured performance, the full red lines the fit of the gain-control model discussed in the paper. Dotted red lines indicate the best fitting Weibull function.*

The normalised total deviance averaged across observers equals 1.58, with all observers ranging between 0.74 and 2.4. This indicates that the quality of fit is fairly good for some observers, but poorer for others. Overall, the gain-control model explains much of the variance in these data. Nevertheless, for three of six observers, total deviance does not belong to the 95% confidence interval of the deviance distribution expected if observers' behaviour were stationary. These confidence intervals were Bonferroni corrected, thus ensuring that the overall probability of making a Type I error, i.e., falsely rejecting the null-hypothesis, equals .05.

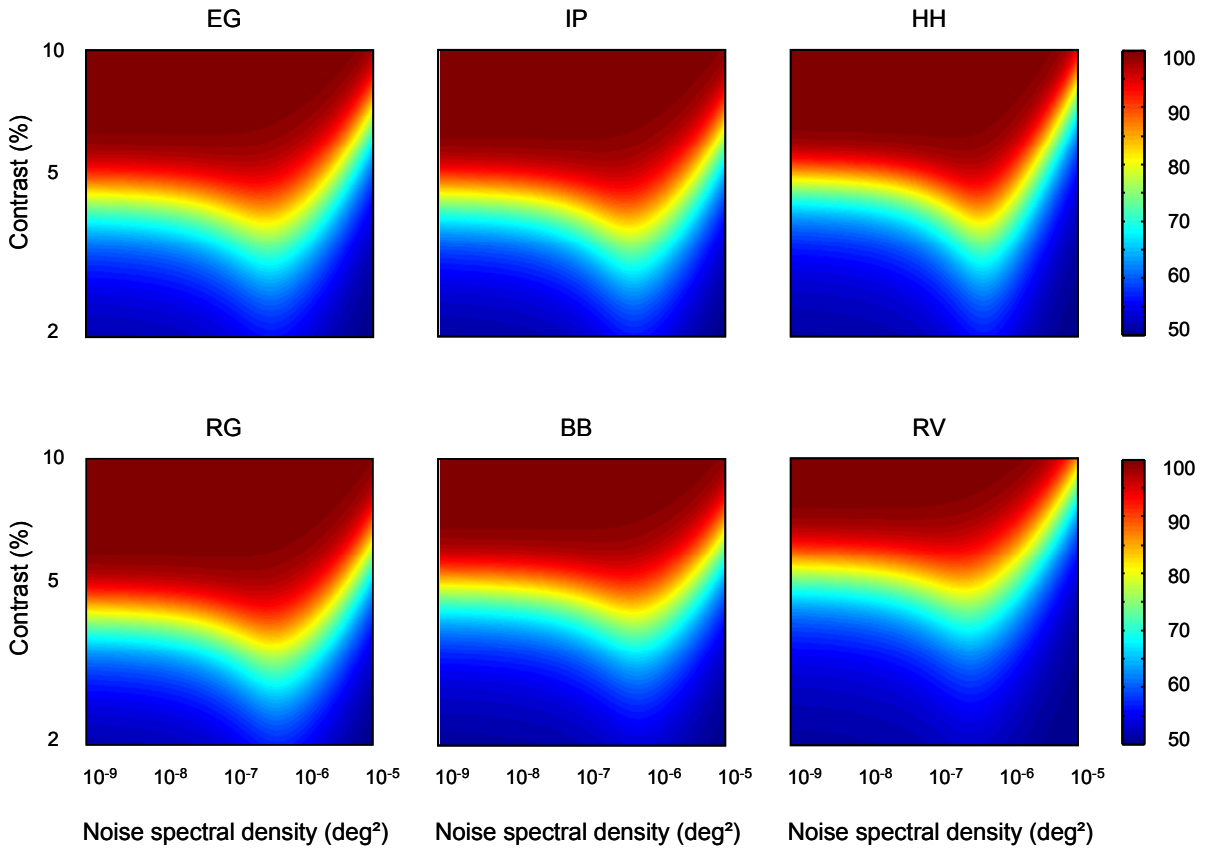

Figure 2-6. *Model fits for all observers. Performance, indicated by colour, is plotted as a function of noise spectral density and signal contrast on double logarithmic coordinates.*

Figure 2-7a shows the model deviance residual distribution across observers, based on 420 blocks of 100 trials each. In Figure 2-7b, model deviance residuals are plotted as a function of noise spectral density with different colour labels for different observers and different symbols for experiment 1 and 2 (squares and circles, respectively). The thick red line describes the mean deviance residual as a function of noise spectral density; the dotted lines indicate the 99% confidence interval. As can be seen in these figures, the distribution of deviance residuals is centred on zero and approximately Gaussian in shape. Because the deviance residuals suggest that the noiseless detection threshold is, perhaps, a bit underestimated and the thresholds in some 'optimal' noise levels a bit overestimated, the model may underestimate the noise facilitation effect slightly. Overall, however, the lack in quality of fit for some observers seems more due to overdispersion (i.e., non-stationary behaviour in some conditions, which cannot be fixed by *any* other model) than to a systematic mismatch.

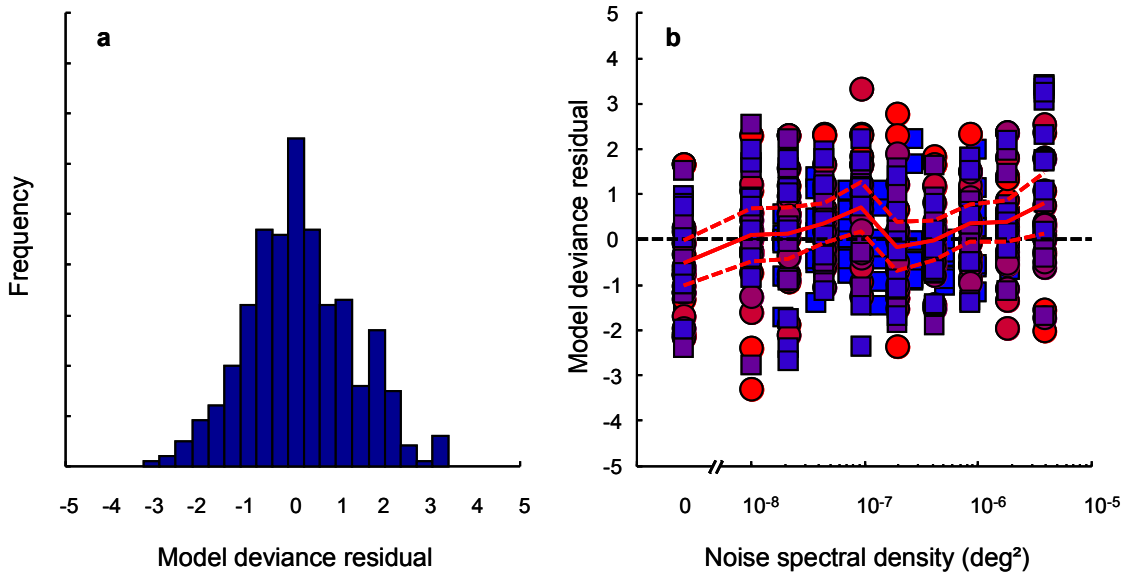

Figure 2-7. ***a.*** *The distribution of model deviance residuals across all observers.* ***b.*** *Model deviance residuals as a function of noise spectral density plotted on semi-logarithmic coordinates. Dotted red lines indicate the 99% confidence interval.*

The parameter estimates listed in table 2-1 are relatively similar across observers. For all participants, the level of early, internal noise (σ^2_{add}) is estimated to be approximately equal to the weakest external noise level that leads to a threshold rise. This is not inconsistent with the notion of 'equivalent input noise' used in some linear detection in noise models (Lu & Dosher, 2008; Nagaraja, 1964).

Sampling (k) is estimated to be approximately one percent, which is in line with some other reported estimates (e.g., Legge, Kersten & Burgess, 1987). Figure 2-8 illustrates the effect of sampling on a signal, a noise and a signal plus noise image. As sampling decreases from 100 % to 10 % (i.e., one log unit), the response of an optimal filter to a signal image decreases by a log unit, while the response to a noise image only decreases by the square root of a log unit (i.e., approximately a factor of 3.16). The response to a signal plus noise image decreases with half a log unit (i.e., approximately a factor of 5) for the noise and contrast levels used in this example. The sampling parameter thus mainly serves to scale the (ratio of) responses to signal, noise and signal plus noise images.

	RV	BB	RG	IP	EG	HH
σ^2_{add}	0.020	0.034	0.026	0.025	0.023	0.012
k	0.007	0.013	0.013	0.012	0.010	0.008
α	3.6e39	4.2e39	8.1e39	3.6e39	1.2e39	4.8e33
β	0.054	0.054	0.054	0.056	0.056	0.14
p	19.9	22.6	22.1	21.46	20.57	13.72
$\boldsymbol{\sigma^2_{late}}$	**1**	**1**	**1**	**1**	**1**	**1**
γ	0.49	0.70	0.45	0.46	1.01	0.07
ξ	6.6e19	4.5e19	7.3e19	5.4e19	2.1e19	1.2e17
λ	0.025	0.022	0.035	0.024	0.053	0.013
D	0.74	2.16*	1.95*	1.50	2.40*	0.74

Table 2-1. *The parameter estimates and deviance of the model described in the text for all observers. Bold symbols and numbers indicate frozen parameter values. * indicates that D is outside the (Bonferroni corrected) 95% confidence interval of a stationary observer.*

Inefficiencies in the visual system need not be conceptualised as sampling. Alternatively, this parameter could be thought of as reflecting the use of a suboptimal filter, for instance a spatial-frequency tuned channel that has an effective bandwidth that is broader than the narrowband Gabor signal to be detected.

As can be seen in equation 2-2 and 2-3, α is simply a rescaling parameter that determines the response range. More interestingly, β reflects the semi-saturation contrast of the contrast response function (Heeger, 1992a,b). For all observers but one (HH), β is estimated to be in the vicinity of the 75% correct (noiseless) detection threshold. The estimates of the response exponent p may seem fairly high compared to fits of the gain-control model to contrast discrimination data (e.g. Foley, 1994; Wichmann, 1999). This is a consequence of the use of external (and early internal) noise: because early noise linearizes nonlinear systems (see discussion) the exponent of the accelerating part of the nonlinearity needs to be high enough to fit the experimentally observed noise benefit.

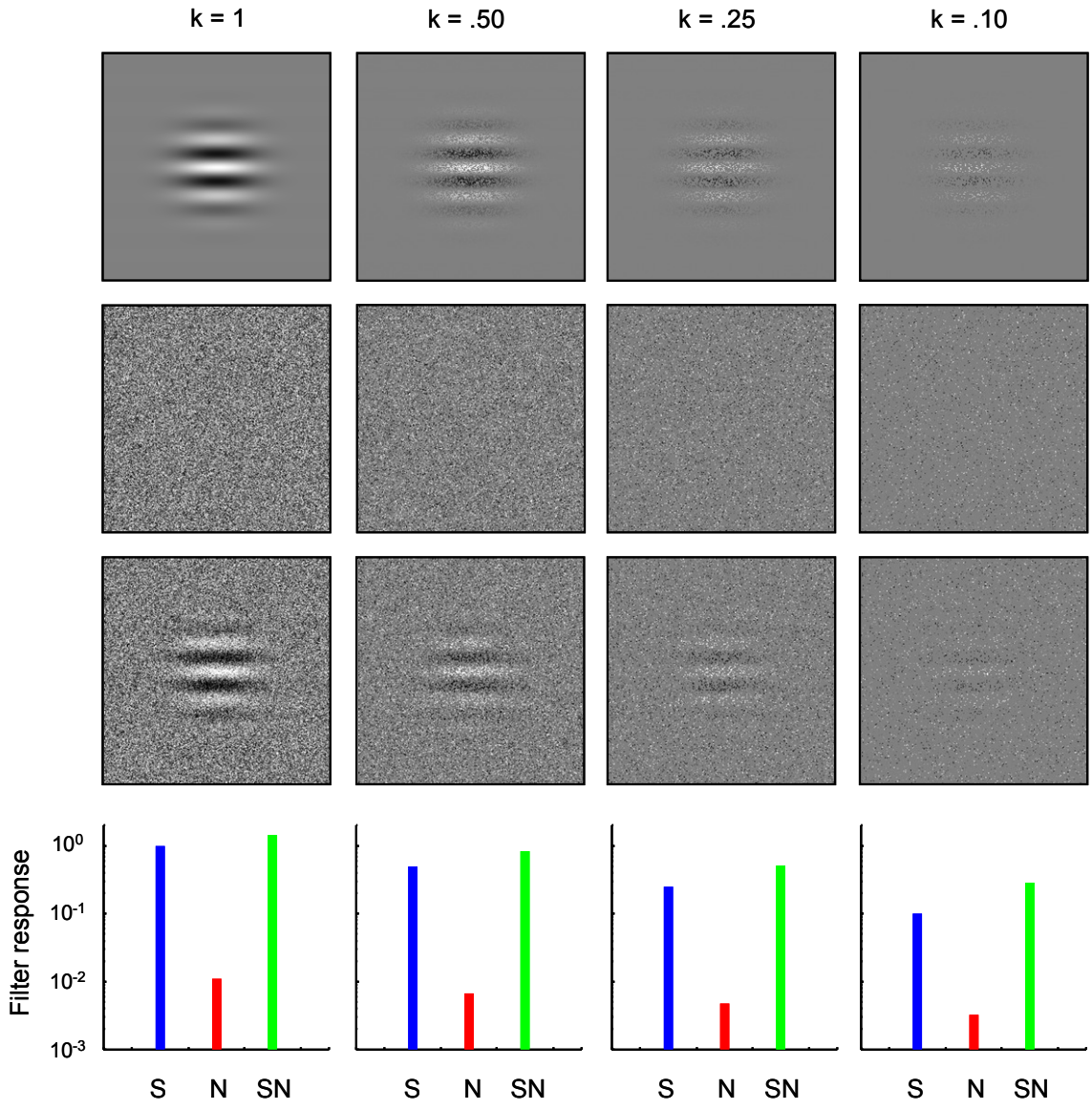

Figure 2-8. *Illustration of the effect of image sampling.* ***Upper row:*** *a maximal contrast stimulus sampled at 100 %, 50 %, 25 % and 10 %.* ***Second row:*** *a noise stimulus sampled at 100 %, 50 %, 25 % and 10 %.* ***Third row:*** *a signal plus noise stimulus sampled at 100 %, 50 %, 25 % and 10 %.* ***Lower row:*** *the (normalised) response of an optimal template to the signal (S), noise (N) and signal plus noise (SN) stimulus. The filter response to the noiseless signal decreases with 1 log unit if sampling decreases with 1 log unit. The response to a noise stimulus decreases with the square root of a log unit, while the response to a signal plus noise stimulus decreases with half a log unit (for these particular noise and contrast levels).*

Inspection of γ reveals that for three of six observers, the exponent of the level-dependent noise source is estimated to be approximately 0.5, which corresponds to a "neural" noise scheme with noise variance proportional to the mean firing rate (e.g.

Geisler & Albrecht, 1997). For one observer, γ is close to one, which corresponds to the standard deviation being proportional to the mean response. With high α-estimates, the proportional constant of the level-dependent noise source, ξ, needs to have high values too to have any influence on the slope of the psychometric functions. Lapse rate (λ) estimates are around two percent, which is not unusual (Wichmann & Hill, 2001a).

In sum, the overall quality of fit provided by the contrast gain-control model is reasonable. While some observers clearly show non-stationary behaviour, there is only a minor indication of systematic errors in the model fit. If any, the depth of the dip is slightly underestimated. Inspection of the parameter estimates reveals that they are not inconsistent with other appplications of the contrast gain-control model in spatial vision.

2.6 Discussion

Much evidence has accumulated that suggests that our early visual system consists of spatial-frequency and orientation-selective channels. In contrast detection circumstances, many data suggest that these channels operate independently and approximately linear, i.e., superposition holds. At higher contrasts, effects of a nonlinear post-filter stage have been reported frequently: Contrast discrimination thresholds initially drop as a function of pedestal contrast, before they rise in a way that is roughly consistent with Weber's law. The standard early vision model explains both types of data as stemming from a linear filter stage, followed by a static nonlinearity and an internal noise source.

It is important to mention that some studies, which made use of slightly more complex stimuli than simple sinusoidal gratings, have reported data that are hard to reconcile with independent spatial-frequency channels (e.g., Derrington & Henning, 1989; Henning, Hertz & Broadbent, 1975). Furthermore, recent work has questioned to what degree the standard model explains the well-known pedestal effect: This effect might be caused by the pooling of information from several channels rather

than stemming from the characteristics of a single spatial-frequency tuned channel (Henning & Wichmann, 2007). Nevertheless, the standard early vision model describes classical contrast discrimination data to a satisfying degree (Wichmann, 1999).

As in the neural analogue of this model, a rectification mechanism at the output of the linear filter stage has been adopted on some occasions as well as here. This rectification mechanism causes external noise to increase the mean filter response to a low contrast stimulus. Because weak filter responses are expanded subsequently, weak amounts of externally added noise could improve contrast detection. This prediction was tested and confirmed in the experiments reported here. Furthermore, both a simple uncertainty reduction mechanism and a global energy discrimination strategy seem to be unlikely explanations. It was also demonstrated that an eight-parameter operationalisation of the contrast gain-control model fitted the data to a reasonable degree, making use of plausible parameter estimates.

At higher noise levels, thresholds rise linearly, as has been reported on many occasions. This linear rise, despite a presumably nonlinear underlying contrast response function, can be understood as a consequence of Birdsall's theorem (Green & Swets, 1966): In a multistage system, sensitive to order, effects of nonlinear transformations occurring after the strongest independent noise injection cannot be measured if this noise source is the crucial performance limiting factor. Thus, once the external noise becomes the crucial performance limiting factor, it effectively linearizes the system. A second implication of this theorem is that a dominant late internal noise source is also a necessary condition for facilitation as a function of noise to occur. Thus, at least in the conditions where facilitation is present, the crucial, performance limiting internal noise source in these experiments is located in the later processing stages, i.e., after the response expansion. Were it located earlier, e.g., in the filter stage, thresholds would increase linearly at all external noise levels.

Finally, it is interesting to note that improved signal transmission as a function of externally added noise is a well-known phenomenon in physics, labelled stochastic

resonance (e.g., Wiesenfeld & Moss, 1995). Stochastic resonance is the signature of nonlinear information processing and has been reported in many man-made and biological systems, including the human tactile (Collins, Imhoff & Grigg, 1996) and auditory system (Zeng, Fu & Morse, 2000). Stochastic resonance is never observed in linear systems, i.e., systems in which superposition holds.

2.7 Conclusions

Addition of weak levels of 2-D, white noise to simple Gabor stimuli, which have typically been used in contrast detection tasks, improves human detection and discrimination performance. This effect is neither caused by reduction of uncertainty about spatial and temporal stimulus occurrence nor is it based on global energy discrimination. We interpret these results as consistent with contrast gain-control models, thus stemming from a rectified linear filter response, followed by response expansion in a post-filter stage and a dominant late internal noise source at low external noise levels. These results imply that detection in noise is better treated as signal discrimination instead of signal detection at the level of a single channel.

2.8 References

Albrecht, D. G., Geisler, W. S., & Crane, A. M. (2003). Nonlinear properties of visual cortex neurons: temporal dynamics, stimulus selectivity, neural performance. In L. Chalupa & J. Werner (Eds.), *The Visual Neurosciences* (pp. 747-764). Boston: MIT Press.

Bethge, M. (2006). Factorial coding of natural images: How effective are linear models in removing higher-order dependencies? *Journal of the Optical Society of America A, 23,* 1253-1268.

Bird, C., Henning, G. B., & Wichmann, F. A. (2002). Contrast discrimination with sinusoidal gratings of different spatial frequency. *Journal of the Optical Society of America A, 19*, 1267-1273.

Blackwell, K. T. (1998). The effect of white and filtered noise on contrast detection thresholds. *Vision Research, 38(2)*, 267-280.

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

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433-436.

Campbell F. W., Carpenter, R. H. S., & Levinson, J. Z. (1969). Visibility of aperiodic patterns compared with that of sinusoidal gratings. *Journal of Physiology (London), 204*, 283-298.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology (London), 197*, 551-556.

Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J, Dan, Y., Olshausen, B. A., Gallant, J. L., & Rust, N. C. (2005). Do we know what the early visual system does? *Journal of Neuroscience, 25(46)*, 10577-10597.

Carandini, M., Heeger, D., & Senn, W. (2002). A synaptic explanation of suppression in visual cortex. *Journal of Neuroscience, 22(22)*, 10053-10065.

Carter, B. E., & Henning, G. B. (1971). The detection of gratings in narrow-band visual noise. *Journal of Physiology (London), 219*, 355-365.

Collins, J. J., Imhoff, T. T., & Grigg, P. (1996). Noise-enhanced tactile sensation. *Nature, 383*, 770.

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America. A, 2(7)*, 1160–1169.

Derrington, A. M., & Henning, G. B. (1989). Some observations on the masking effect of two-dimensional stimuli. *Vision Research, 29(2)*, 241-246.

DeValois, R. L., & DeValois, K. K. (1988). *Spatial Vision.* Oxford University Press.

Dosher, B. A. & Lu, Z.-L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences, 95(23)*, 13988-13993.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: masking experiments require a new model. *Journal of the Optical Society of America A, 11(6)*, 1710–1719.

Foley, J. M., & Chen C.C. (1997). Analysis of the effect of pattern adaptation on pattern pedestal effects: A two-process model. *Vision Research, 37*, 2779-2788.

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

Freeman, T. C. B., Durand, S., Kiper, D. C., & Carandini, M. (2002). Suppresion without inhibition in visual cortex. *Neuron, 35*, 759-771.

Geisler, W. S., & Albrecht, D. G. (1995). Bayesian analysis of identification performance in monkey visual cortex: Nonlinear mechanisms and stimulus certainty. *Vision Research, 35(19)*, 2723-2730.

Geisler, W. S., & Albrecht, D. G. (1997). Visual cortex neurons in monkeys and cats: Detection, discrimination, and identification. *Visual Neuroscience, 14(5)*, 897-919.

Gorea, A., & Sagi, D. (2001). Disentangling signal from noise in visual contrast discrimination. *Nature Neuroscience, 4(11)*, 1146-1150.

Goris, R. L. T., Wagemans, J., & Wichmann, F. A. W. (2008). Modelling contrast discrimination data suggests both the pedestal effect and stochastic resonance to be caused by the same mechanism. *Journal of Vision, 8(15):17, 1-21.*

Graham, N. (1989). *Visual Pattern Analyzers*. Oxford University Press.

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: a comparison of single channel and multichannel models. *Vision Research, 11*, 251-259.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics.* New York: John Wiley & Sons, Inc.

Heeger, D. (1992a). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9*, 181-197.

Heeger, D. (1992b). Half-squaring in responses of cat striate cells. *Visual Neuroscience, 9*, 427-443.

Henning, G. B. (1988). Spatial-frequency tuning as a function of temporal frequency and stimulus motion. *Journal of the Optical Society of America A, 5*, 1362-1373.

Henning, G. B., Bird, C. M., & Wichmann, F. A. (2002). Contrast discrimination with pulse trains in pink noise, *Journal of the Optical Society of America A, 19*, 1259-1266.

Henning, G. B., Derrington, A. M., & Madden, B. C. (1983). Detectability of several ideal spatial patterns. *Journal of the Optical Society of America, 73*, 851-854.

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

Henning, G. B., Hertz, B. G., & Hinton, J. L. (1981). Effects of different hypothetical detection mechanisms on the shape of spatial-frequency filters inferred from masking experiments: I. noise masks. *Journal of the Optical Society of America, 71*, 574-581.

Henning, G. B., & Wichmann, F. A. (2007). Some observations on the pedestal effect. *Journal of Vision, 7(1:3)*, 1-15.

Holmes, D. J. & Meese, T. S. (2004). Grating and plaid masks indicate linear summation in a contrast gain pool. *Journal of Vision*, *4(12:7)*, 1080-1089.

Kontsevich, L. L., Chen, C.-C., & Tyler, C. W. (2002). Separating the effects of response nonlinearity and internal noise psychophysically. *Vision Research, 42*, 1771-1784.

Legge, G. E. (1981). A power law for contrast discrimination. *Vision Research, 21(4)*, 457-467.

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

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

Lu, Z.-L., & Dosher, B. A. (2008). Characterizing observers using external noise and observer models: assessing internal representations with external noise. *Psychological review, 115(1),* 44-82.

Nachmias, J. (1989). Contrast modulated maskers: test of a late nonlinearity hypothesis, *Vision Research* 29 (1), 137–142.

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

Nagaraja, N.S. (1964). Effect of luminance noise on contrast thresholds, *Journal of the Optical Society of America 54(7)*, 950-955.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America. A, 2(9)*, 1508-1532.

Pelli, D. G. (1991). Noise in the visual system may be early. In: M. Landy & J. A. Movshon (Eds.), *Computational Models of Visual Processing* (pp. 147-152). Cambridge: MIT Press.

Pelli, D. G. (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*, 437-42.

Pelli, D. G., & Farell, B. (1999). Why use noise? *Journal of the Optical Society of America A, 16*, 647-653.

Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nature Neuroscience, 4(8),* 819-825.

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

Stromeyer, C.F., & Julesz, B. (1972). Spatial frequency masking in vision: Critical bands and spread of masking, *Journal of the Optical Society of America 62*, 1221–1232.

Watson, A. B., Barlow, H. B., & Robson, J. G. (1983). What does the eye see best? *Nature 302,* 419-422.

Wichmann, F. A. (1999). *Some aspects of modelling human spatial vision: Contrast discrimination*. Unpublished doctoral dissertation, The University of Oxford, Oxford, UK.

Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness-of-fit. *Perception & Psychophysics, 63*, 1293-1313.

Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics, 63*, 1314-1329.

Wilson, H. R., McFarlane, D. K., & Philips, G.C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research 23(9),* 873-882.

Wiesenfeld, K., & Moss, F. (1995). Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDS. *Nature, 373*, 33-36.

Yang, J., & Makous, J. (1995). Modeling pedestal experiments with amplitude instead of contrast. *Vision Research, 35(14)*, 1979-1989.

Zeng, F.-G., Fu, Q.-J., & Morse, R. P. (2000). Human hearing enhanced by noise. *Brain Research, 869,* 251-255.

Chapter 3 Modelling contrast discrimination data suggests both the pedestal effect and stochastic resonance to be caused by the same mechanism.

Computational models of spatial vision typically make use of a (rectified) linear filter, a non-linearity and dominant late noise to account for human contrast discrimination data. Linear-non-linear cascade models predict an improvement in observers' contrast detection performance when low, sub-threshold levels of external noise are added (i.e., stochastic resonance). Here, we address the issue whether a single contrast gain-control model of early spatial vision can account for both the pedestal effect, i.e., the improved detectability of a grating in the presence of a low-contrast masking grating, and stochastic resonance. We measured contrast discrimination performance without noise and in both weak and moderate levels of noise. Making use of a full quantitative description of our data with few parameters combined with comprehensive model selection assessments, we show the pedestal effect to be more reduced in the presence of weak noise than in moderate noise. This reduction rules out independent, additive sources of performance improvement and, together with a simulation study, supports the parsimonious explanation that a single mechanism underlies the pedestal effect and stochastic resonance in contrast perception.

This chapter has been published as:

Goris, R. L. T., Wagemans, J., & Wichmann, F. A. W. (2008). Modelling contrast discrimination data suggests both the pedestal effect and stochastic resonance to be caused by the same mechanism. *Journal of Vision, 8(15):17, 1-21.*

The paper can be downloaded from: http://journalofvision.org/8/15/17/

3.1 Introduction

Threshold detection studies have provided much evidence consistent with a model of our visual system in which visual information is analysed in relatively independent and approximately linear spatial-frequency and orientation-selective filters (Blakemore & Campbell, 1969; Campbell, Carpenter & Levinson, 1969; Campbell & Robson, 1968; DeValois & DeValois, 1988; Graham & Nachmias, 1971). The corresponding spatial weighting functions of such band-limited channels are slightly asymmetric on double logarithmic coordinates (Henning, 1988; Henning, Hertz & Hinton, 1981), but are nonetheless typically approximated by Gabor functions. Though not an "optimal" stimulus (Watson, Barlow & Robson, 1983; but see Henning, Derrington & Madden, 1983), Gabors are popular stimuli to explore contrast detection because they are band-limited in spatial-frequency and localised in space (Daugman, 1985). Studies that made use of more complex stimuli than single component gratings have reported results that are difficult to reconcile with the aforementioned independence and linearity (Derrington & Henning, 1989; Henning, Bird & Wichmann, 2002; Henning, Hertz & Broadbent, 1975), but the multi-channel model still captures much of what we understand of early spatial vision.

To gain insight in visual processing at supra-threshold contrasts – a prerequisite for any model of spatial vision to be truly useful – sinusoidal contrast discrimination has been studied extensively (e.g., Baker, Meese & Georgeson, 2007; Bird, Henning & Wichmann, 2002; Foley, 1994; Foley & Chen, 1997; Foley & Legge, 1981; Georgeson & Georgeson, 1987; Gorea & Sagi, 2001; Henning & Wichmann, 2007; Kontsevich, Chen & Tyler, 2002; Legge, 1981; Legge & Foley, 1980; Legge, Kersten & Burgess, 1987; Nachmias & Sansbury, 1974; Wichmann, 1999; Yang & Makous, 1995). The main finding of these studies is the so-called pedestal effect, i.e., the improved detectability of a grating in the presence of a low-contrast "pedestal" stimulus of identical spatial frequency, phase and orientation. At higher pedestal contrasts, discrimination thresholds rise in a way that is roughly consistent with Weber's law. Analogues of the pedestal effect in spatial vision have been reported in other sensory modalities such as amplitude discrimination in hearing and, recently, flicker discrimination in vision (Smithson, Henning, MacLeod & Stockman, under review).

The mechanisms suggested to underlie the dipper-shaped thresholds-vs.-contrast (TvC) function in spatial vision include non-linear transduction (e.g., Legge & Foley, 1980; Yang & Makous, 1995), contrast gain-control (e.g., Foley, 1994; Foley & Chen, 1997) and uncertainty reduction (e.g., Pelli, 1985). While the first two types of models – indistinguishable for single component sinusoidal contrast discrimination (Wichmann, 1999) – effectively make use of an expansive non-linearity in a sensory post-filter stage to explain the pedestal effect, uncertainty models leave the sensory processing linear but place the non-linearity, required to explain the dipper function, in the decision process. In effect, all models that explain the dipper-function with some degree of success make use of a (rectified) linear filter stage, a non-linear post-filter stage and a dominant late internal noise source.

Recently, Goris, Zaenen and Wagemans (2008) noted that such linear-non-linear cascade models predict an improvement in observers' performance when low, sub-threshold levels of external noise are added. The improvement arises because rectification causes the external noise to increase the mean filter response to both a weak signal-plus-noise and a noise stimulus. Because these responses are expanded subsequently, the difference between the means of the internal response distributions representing the noise and signal-plus-noise increases. If the crucial performance-limiting noise source is located in later processing stages, this increased difference between the means will lead to a better signal-to-noise ratio at the decision stage, resulting in improved contrast detection.

Consistent with this prediction, Goris et al. showed that detection thresholds reach a minimum for very low noise levels. This type of phenomenon – improved signal transmission due to the addition of external noise – is termed stochastic resonance in physics (Wiesenfeld & Moss, 1995). Similar observations have been reported once before in spatial vision (Blackwell, 1998), and in other sensory modalities such as tactile (Collins, Imhoff & Grigg, 1996) and auditory perception (Zeng, Fu & Morse, 2000). To test their speculation that the mechanism underlying the pedestal effect also leads to stochastic resonance, Goris et al. fitted a contrast gain-control model to their detection-in-noise data and found the fits to be reasonable.

Aiming to explain the pedestal effect, the Goris et al. gain-control model assumes that the effect characterises the operation of individual, spatial-frequency tuned channels. Recently, it has been suggested that this is not the case in spatial vision (Henning & Wichmann, 2007). The disappearance of the pedestal effect in notched noise led Henning and Wichmann (2007) to argue that the pedestal effect may be characteristic of the way in which observers use information from the spatial-frequency and orientation-selective channels tuned to frequencies and orientations away from the signal rather than characteristic of a single channel tuned to the spatial characteristics of the signal to be detected. The gain-control model may thus not be an *explanation* of single channel behaviour despite offering an excellent *description* of contrast discrimination performance (e.g., Wichmann, 1999). In this paper, we exploit the gain-control model as a powerful statistical tool with few free parameters to describe our contrast discrimination data, and we use the fits to different data sets to make statistically sound inferences about changes in the data.

One way of observing effects of experimental manipulations on contrast discrimination is to normalise all pedestal and signal contrasts by the detection threshold measured in the absence of a pedestal. When this is done, data are expressed on the same scale, irrespective of differences in detection threshold. Once corrected for absolute visibility, the dipper-shaped TvC-function has been reported to be remarkably invariant to parameter manipulations such as retinal illumination (e.g., Yang & Makous, 1995), retinal eccentricity (e.g., Bradley & Ohzawa, 1986), varying degrees of stimulus uncertainty (e.g., Foley & Schwartz, 1998) and cross-surround facilitation (Yu, Klein & Levi, 2002).

Measurements in the presence of broadband noise have revealed a similar invariance (e.g., Pelli, 1985), though the depth of the dip is reduced for at least some observers (e.g., Figure 3-7 in Henning & Wichmann, 2007). This might be a consequence of the additional stimulus variability that is introduced: In the presence of noise the pedestal stimulus might have an *average* optimal contrast, but on each trial, due to the noise, the contrast in the relevant channel will be somewhat higher or lower and thus not optimal. This will lead to a higher threshold, i.e., a reduced pedestal effect.

Considered this way, the invariance to the presence of noise displayed by some observers, especially in 1-D noise (e.g., Figure 3-8 and 10 in Henning & Wichmann, 2007), may indicate a change in the underlying mechanism.

In 2-D noise, the number of active channels, sensitive to the signal, is similarly important: Having few active channels, perhaps even with correlated noise (Henning et al., 2002) could be expected to reduce the size of the pedestal effect due to the additionally introduced stimulus variability. On the other hand, having many active channels that sample different image regions and hence different regions of noise could enable the averaging out of noise-induced effects of signal variability and thus leave the size of the pedestal effect (almost) unchanged. Whatever the exact mechanism underlying contrast discrimination in noise may be, our main interest here concerns the dipper-effect being approximately invariant to addition of (strong) broadband noise.

In sum, all psychophysical spatial vision models that aim to explain the pedestal effect make use of a (rectified) linear filter stage, a non-linear post-filter stage and a dominant late internal noise source. This implies that these models actually predict that stochastic resonance occurs in contrast detection. This prediction has been confirmed experimentally (Goris et al., 2008). Here, we address the issue whether a single model of early spatial vision, having a (rectified) linear filter, a non-linearity and dominant late noise can account for both the pedestal effect *and* stochastic resonance.

If the same mechanism underlies the pedestal effect and stochastic resonance, we should expect the pedestal effect to be reduced in the presence of weak noise. The logic of our argument is as follows: The mean response of a linear filter to a low contrast signal is raised in the presence of external noise due to rectification. As explained above, for certain external noise levels, the subsequent non-linear stage leads to an improved discriminability of the signal in the zero-pedestal contrast case. However, as we shall see later, as pedestal contrast increases from zero, the response-enhancing effect of rectification diminishes and soon disappears completely. From

these pedestal contrasts on, the only effect of weak noise is that it introduces additional variability. Thus, at an 'optimal' noiseless pedestal contrast, addition of external noise will not lead to a benefit but rather hurt contrast discrimination performance. Therefore, if one mechanism underlies both facilitation phenomena, addition of weak noise can only improve discrimination performance at pedestal contrasts lower than the 'optimal' pedestal contrast, implying that the size of the pedestal effect is reduced in the presence of weak noise.

An alternative hypothesis may be to argue that such a reduced dipper effect in weak noise is a consequence of the stimulus variability introduced by the noise (see above). In this view, strong noise (which produces even more stimulus variability) should lead to a stronger reduction in the dipper effect than weak noise does. As our data will bear out, this is not the case, however: the reduced dip in weak noise is not simply a consequence of increased stimulus variability.

3.2 Methods

3.2.1 Equipment

The experiments were run on a Power Macintosh G4 computer using the software packages MATLAB (MathWorks, Natick, MA) and PSYCHTOOLBOX (Brainard, 1997; Pelli, 1997a). Gamma correction using an 8-bit lookup table ensured that the monitor was linear over the entire luminance range used in the experiments. The stimuli were presented on a SONY Trinitron GDM-FW900 monitor with a spatial resolution of 1920 × 1440 pixels and a temporal resolution of 75 Hz. The experiment was run in a darkened room and the screen's mean grey background luminance was 42 cd/m^2. Viewing distance was 120 cm, leading to a pixel-size of .009° of visual angle.

3.2.2 Observers

Four observers participated (L.V., E.G., B.B., and L.V.E.). All were well practiced in the task and familiar with the stimuli before data collection began and had normal or

corrected-to-normal vision. All observers were naïve to the purpose of the experiment.

3.2.3 Stimuli

The Gabor stimuli consisted of horizontally orientated sine-gratings with a spatial frequency of 7 c/deg, which were multiplied by a circularly-symmetric two-dimensional spatial Gaussian envelope with a σ of 0.27°. Stimuli had a spatial extent of 2.35° of visual angle.

For each stimulus presentation in the noise conditions a fresh noise sample was generated at every pixel, sampled from independent identically distributed Gaussian distributions centred at the mean luminance. For each observer – in addition to the no-noise condition – two noise conditions were run; for the noises noise-power spectral densities ranging between 0.84 and 42 $\times$ 10^{-7} deg^2 were used. Noise-power spectral density is defined as the luminance variance multiplied by the pixel area, expressed in visual degrees squared. Effects of temporal waveform and duration are not considered here. It is proportional to the average power at the different frequencies present in the noise. The maximal amount of clipping (i.e., pixels set to the minimal or maximal luminance values because of the limited 8-bit dynamic range of the DACs on the video card) at the highest noise level was around 2.5%. Through simulations in MATLAB this level of clipping was calculated to have no significant influence on the spectral properties of the Gaussian white noise.

3.2.4 Procedure

The same procedure as described in Goris et al. (2008) was adopted. A temporal two-alternative-forced-choice (2AFC) task was used. Stimulus presentation time was approximately 12.5 ms for subject B.B., 25 ms for subject L.V.E. and 50 ms for subject E.G. and L.V.. We used short presentation times to obtain detection thresholds around 5% Michelson contrast for each observer – due to their different absolute sensitivities we thus had to use different presentation times for our observers. Signal presentation time has a significant influence on the parameters of the linear-nonlinear model typically used to fit the TvC functions (Wichmann, 1999). However, as all our analyses will be performed on the data of individual observers,

differences resulting from the different presentation times are immaterial with respect to the conclusions drawn.

We attempted to keep detection thresholds comparatively high at 5% contrast to reduce the potential influence of artefacts due to the non-linear monitor operations (i.e., power saving features, gamma correction, luminance rounding etc.; Bach, 1997; Bach, Meigen & Strasburger, 1997; Brainard, 1989; Naiman & Makous, 1992; Pelli, 1997b; Wichmann, 1999). Estimates of the effective images ensured that these non-linear monitor operations did not significantly affect the spectral properties of the noise stimuli. To minimize potential effects of spatial and temporal stimulus uncertainty, the stimulus was surrounded by a red square (2.35°) at full contrast; this square appeared and disappeared with the stimulus onset and offset.

All trials started with the 266-ms presentation of a grey field of mean luminance. The first stimulus presentation was followed by an inter-stimulus-interval (ISI) of 466 ms, and then followed by the second stimulus presentation. The ISI was thus about 20 times longer than the stimulus duration and more than three times as long as current estimates of the temporal impulse response of human spatial frequency channels (Graham, 1989). The noise was turned on and off with both stimulus presentations. Independent noise samples were presented in the two observation intervals. Response time was limited to 1,000 ms, and was indicated by a grey field of mean luminance, presented after the second stimulus presentation. After the response screen, a screen with a green square (2.35°) appeared for about 750 ms. Observers received auditory feedback after each trial; missed trials (about 1%) were shown again; on these occasions the interval containing the signal stimulus was selected randomly again and the noise was drawn anew. After a few training blocks participants were familiar with this steady temporal task profile.

Before data collection began, all observers participated in a detection-in-noise task that consisted of 2,000 trials. These training data were used to estimate the noise level that would evoke maximal facilitation – stochastic resonance – as a function of noise power density for each observer (i.e., the noise level used in the weak-noise

condition, which varied somewhat across observers). Across observers, the weak noise had an average power spectral density of 3.9×10^{-7} deg^2; and the moderate noise of 42×10^{-7} deg^2, approximately a log unit increase. We call the highest noise level used in this experiment 'moderate' because it leads, on average, only to a rise by a factor of 1.34 in the detection threshold at 75% correct. The high detection thresholds purposely enforced through the use of very short presentation times, together with the limited dynamic range of today's graphics cards and CRTs which prevented us from using higher noise power densities, unfortunately prevented us from having stronger masking effects.

Ten pedestal contrast levels, ranging between 0 and 40% were used. To obtain psychometric functions, six increment contrasts were tested at each noise and pedestal level. Within each block of 50 trials, 50 conditions were randomly sampled from all 180 conditions (3 noise power densities * 10 pedestal contrasts * 6 increment contrasts). To help observers to maintain constant eye accommodation, a red square of 3.5° of visual angle, surrounding the stimuli was presented throughout the experiment. Subjects L.V. and B.B. completed 9,000 trials in total. Because one data point heavily influenced the initial estimate of subject E.G.'s noiseless detection threshold, she completed 700 additional detection trials and thus 9,700 trials in total. Subject L.V.E. completed only 4,500 trials in total – 25 instead of 50 trials per condition – and no ethically acceptable means could be found to motivate her to finish the experiment.

3.3 Results

A summary of the results for all observers is shown in Figure 3-1. Figure 3-1a displays the data of all observers in the three detection conditions, i.e., detection without noise and detection in the presence of weak and moderate noise. As for all figures in this paper, red symbols and lines refer to the no-noise condition, green to the weak-noise condition and blue to the moderate-noise condition. Different symbols refer to different observers. To express the data of all observers on the same scale, the contrasts of the stimuli of each noise condition were first normalised by the 75%-

correct detection threshold of each observer in that condition and then multiplied by the average detection threshold of all observers in that condition. The psychometric functions relating the percentage of correct responses to the logarithm of signal contrast were fit with Weibull functions using the maximum-likelihood procedure of Wichmann and Hill (2001a,b); 100,000 bootstraps were run to estimate confidence intervals.

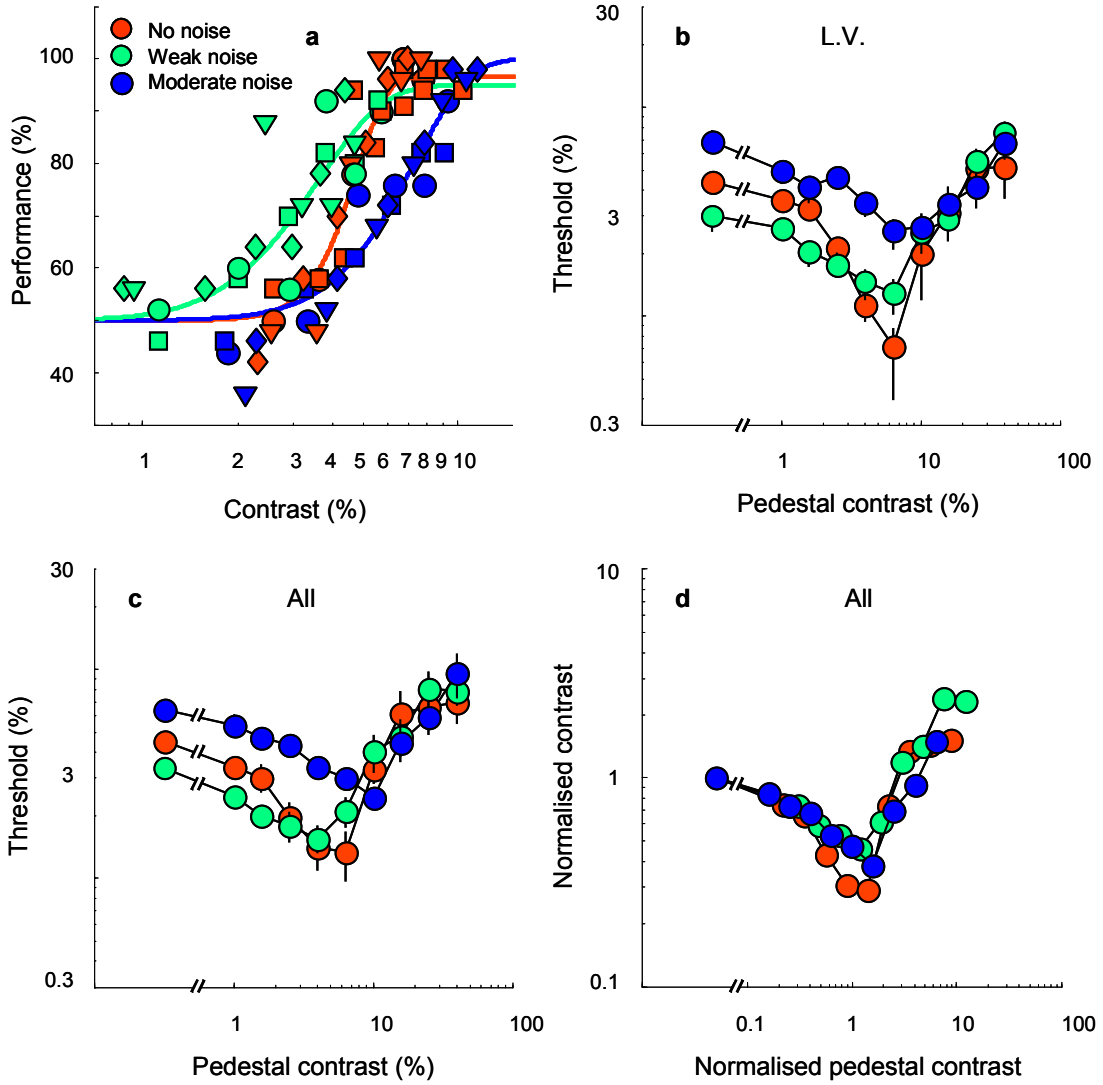

Figure 3-1. ***a.*** *Performance as a function of signal contrast on semi-logarithmic coordinates for the three detection conditions. Red symbols refer to the no-noise condition, green to the weak-noise condition and blue to the moderate-noise condition. Full lines indicate the best fitting Weibull-functions to the pooled data of all observers (L.V.: ●, E.G.: ■, B.B.: ♦ and L.V.E.: ▼).* ***b.*** *The complete TvC-functions at 75% correct for observer L.V..* ***c.*** *The complete TvC-functions at 75% correct averaged over observers.* ***d.*** *The same functions as in* ***c****, normalised by their detection thresholds.*

As can be seen in Figure 3-1a, we find stochastic resonance as in Goris et al.: Addition of weak noise improves contrast detection performance. This is not the case for moderate noise, which increased the 75%-correct threshold by a factor of 1.34. Psychometric functions measured with and without weak noise added to the stimuli are not parallel on semi-logarithmic coordinates. Even though the weak noise is barely visible, the slope of the psychometric function is shallower than in the noiseless detection condition. For the moderate-noise condition, the estimated slope is in-between (and does not differ significantly from the two other conditions given the amount of data collected).

This change in slope as a function of noise level is consistent with earlier observations on the slope of psychometric functions in contrast discrimination: Psychometric functions are steepest without any pedestal added and most shallow in the trough of the dip (Nachmias & Sansbury, 1974; Wichmann, 1999). As in contrast discrimination, this implies that the strength of the noise facilitation effect depends on the performance level considered, i.e., the size of the threshold reduction is inversely related to performance level. For our current data the facilitation at 60% correct is estimated to be of a factor 1.74 (i.e., the ratio of the 60% correct thresholds measured without noise and with weak noise added equals 1.74), at 75% correct this factor is 1.39, and at 90% it is 1.15. Though the particular facilitation factors did vary somewhat over observers (ranging between 1.26 and 1.47 at 75% correct), all individual datasets displayed the trends described above.

We now consider the results at all pedestal contrasts. Figure 3-1c shows the full TvC-functions at 75% correct for all noise conditions. The most leftward points denote detection (i.e., the pedestal contrast equals zero). For illustration purposes, thresholds were simply averaged over observers, without any attempt to rescale the data. The normalised versions of these curves can be seen in Figure 3-1d, in which threshold and pedestal contrasts were normalised by the average test contrast threshold in each of the three noise conditions (when the pedestal was zero). Despite the simple averaging of the data, the TvC-functions capture the main trends that can be seen in all individual data-sets (one example is shown in Figure 3-1b).

Figures 3-1(b-c) illustrate that at the lowest pedestal levels the data shown in green lie below the no-noise data, shown in red. For low pedestal contrasts, we thus find stochastic resonance, i.e., contrast discrimination performance benefits from the presence of weak noise. At slightly higher pedestal contrasts, this effect disappears and eventually, around the trough of the dip for the no-noise condition, weak noise hinders performance. We thus find no 'super-dipper' in the presence of weak noise, as for instance has been reported for contrast discrimination in combination with cross-surround facilitation (Yu et al., 2002). Stochastic resonance and the pedestal effect are not simply additive and thus not two independent effects, in line with our hypothesis that a single mechanism accounts for both.

Furthermore, it can be seen in the normalised plots in Figure 3-1d that weak noise seems to reduce the pedestal effect to a similar, perhaps even stronger degree as moderate noise does. For these averaged data, maximal threshold reduction at 75% correct was of a factor 3.44 without noise (ranging between 3.15 and 6.13 over observers), 2.17 in the presence of weak noise (ranging between 1.96 and 3.53) and 2.62 in the presence of moderate noise (ranging between 2.28 and 2.99). A stronger reduction of the dipper effect in weak noise than in strong noise is not consistent with the hypothesis that this reduction is solely due to the introduced stimulus stochasticity. However, for all observers, the 95% confidence intervals of the maximal threshold reduction factors of all conditions overlapped, pointing to the need to use more sophisticated statistical techniques to analyse our data (see the modelling below).

In sum, we find stochastic resonance, i.e., a significantly reduced detection threshold in the presence of weak noise. Furthermore, when data of all observers are pooled, the slope of the psychometric function is significantly decreased in the presence of weak noise. The full TvC-functions at 75% correct suggest that the pedestal effect is diminished in the presence of both weak and moderate noise, though this reduction is not significant if assessed for each observer independently. The failure to find a 'super-dipper' at 75% correct suggests that stochastic resonance and the pedestal effect are *not* independent effects.

Of course, only considering 75%-correct-thresholds ignores much of the information present in our data-set. To study whether and how the presence of noise changes contrast discrimination performance in more detail, we will make use of the standard gain-control model introduced by Foley (1994) and elaborated by Wichmann (1999). In this application, we mainly use this model as a statistical tool to get a full quantitative description of our data with as few parameters as possible.

3.4 Model: Equations, Fitting, Evaluation and Selection

3.4.1 Equations

We used the standard divisive gain-control model as formalised by Wichmann (1999). Observer's responses are modelled within the framework of signal detection theory (SDT, Green & Swets, 1966). Three hypothetical stages are specified in SDT-models: First, a *stimulus theory* describes how a transduction mechanism maps physical stimuli to internal states, second, a probabilistic *theory of internal states* describes the probability distribution of the internal states that results from repeated presentation of the same stimulus, and finally a deterministic *response theory* describes a decision rule that maps internal states to a response.

In the gain-control model, the transduction mechanism was chosen to be the generalized four parameter Naka-Rushton function (free parameters α, β, η, and κ). These parameters express the response gain (α), the semisaturation contrast (β), the response exponent (η) and the gain-control exponent (κ) of the contrast response function. One additional free parameter was added to describe the internal noise, assumed to be Gaussian and signal-independent (free parameter σ).

Recently, there has been some debate in spatial vision regarding the question whether a signal-dependent source might also contribute to internal noise (e.g., Georgeson & Meese, 2006; Gorea & Sagi, 2001; Kontsevich et al., 2002; Wichmann, 1999). Level-dependent noise may be needed to explain contrast discrimination performance at high pedestal levels. Here it was not needed, because the majority of data points were gathered at relatively low pedestal levels. Thus we did not include level-dependent

noise for our modelling to reduce the number of free parameters. This is not to argue that level-dependent noise may not be crucial; indeed, one of us showed that level-dependent noise is critically needed to fit contrast discrimination data at high pedestal contrasts (Wichmann, 1999).

As is standard, it was assumed that the observer's response ('interval 1' or 'interval 2') is determined by the stimulus interval that led to the highest internal state. The model equations (see equation 3-1–3-3) were arranged to express percent correct, $p(\Delta x, x)$, as a function of the contrast increment (Δx) and the pedestal contrast (x) in a 2AFC-task.

$$p(\Delta x, x) = \int_0^\infty \frac{1}{\sqrt{2\pi g(\Delta x, x)}} e^{-\frac{(z - f(\Delta x, x))^2}{2g(\Delta x, x)}} dz\,, \tag{3-1}$$

where z is a dummy variable, and $f(\Delta x, x)$ and $g(\Delta x, x)$ are given by

$$f(\Delta x, x) = \alpha\left(\frac{(\Delta x + x)^\eta}{\beta^\kappa + (\Delta x + x)^\kappa} - \frac{x^\eta}{\beta^\kappa + x^\kappa}\right), \tag{3-2}$$

and

$$g(\Delta x, x) = 2\sigma^2\,. \tag{3-3}$$

3.4.2 Fitting

One of the five parameters can be arbitrarily set to any value. To follow the usual convention, σ was taken to be 1, resulting in a four free parameter model. An additional (highly constrained) vector of free parameters $\boldsymbol{\lambda}$ ("lapse rates"), estimated for each noise and pedestal combination as explained by Wichmann and Hill (2001a,b) was introduced in the fitting of the model to avoid biased parameter estimates (for details see Wichmann, 1999).

Priors were introduced for each parameter to constrain estimates to realistic values. To find the surface $p(\Delta x,\ x)$ that maximizes the likelihood that the data were generated from a process with success probability given by $p(\Delta x,\ x)$, the log-likelihood of the surface $p(\Delta x,\ x)$ given the parameters (α, β, η and κ) was maximized using purpose-written software in MATLAB (*fminsearch*, which makes use of the Nelder-Mead simplex search method). The log-likelihood of the surface $p(\Delta x,\ x)$ given parameter vector $\boldsymbol{\theta}$, containing {α, β, η, κ} with σ = 1 and $\boldsymbol{\lambda}$ equal to the lapse rate vector derived from the psychometric function fits, is given by equation 3-4:

$$l(\theta) = \sum_{j=1}^{Z}\sum_{i=1}^{K_j} \log\binom{n_{ji}}{y_{ji}n_{ji}} + y_{ji}n_{ji}\log\left(p\left(\Delta x_{ji}, x_j; \theta\right)\right) + \left(1 - y_{ji}\right)n_{ji}\log\left(1 - p\left(\Delta x_{ji}, x_j; \theta\right)\right), \qquad (3\text{-}4)$$

with n_{ji} the number of trials (block size) measured at pedestal contrast j and signal contrast i and y_{ji} the proportion of correct responses in that condition. Because the problem is non-convex due to $\boldsymbol{\lambda}$, a multi-start procedure with semi-randomly chosen initial parameter values was used. For each model fit reported, at least 20 different starting points were used.

3.4.3 Evaluation

To evaluate model fits we considered the overall distance between model prediction and data and the presence of systematic errors in the residuals. Quality of the overall fit can be assessed by judging total deviance (see equation 3-5), i.e., the log-likelihood ratio of the saturated model and the best fitting model (the saturated model is the model with no residual error between model predictions and data). What deviance does not assess, however, are systematic trends in the deviance residuals (see equation 3-6), i.e., the agreement between individual data points and the corresponding model prediction. For binomial data, deviance is expressed by equation 3-5.

$$D = 2\sum_{j=1}^{Z}\sum_{i=1}^{K_j}\left\{n_{ji}y_{ji}\log\left(\frac{y_{ji}}{p\left(\Delta x_{ji}, x_j\right)}\right) + n_{ji}\left(1 - y_{ji}\right)\log\left(\frac{1 - y_{ji}}{1 - p\left(\Delta x_{ji}, x_j\right)}\right)\right\}, \qquad (3\text{-}5)$$

This statistic indicates how well a model describes data. Asymptotically, it can be shown to be χ^2-distributed, with degrees-of-freedom equal to the number of data blocks minus the number of free parameters if the model is correct and an observer behaves perfectly stationary during the whole experiment (and thus generates truly binomially distributed data). Often, due to a variety of reasons, this is not the case. Responses of non-stationary observers are more variable than binomially distributed data and thus lead to higher deviances (overdispersion).

Wichmann (1999) has shown that, due to the typically relatively small number of measurements, the asymptotically derived deviance distributions often fail to approximate the real deviance distribution for psychophysical data-sets. The real deviance distribution can be estimated easily by means of Monte Carlo simulations. As suggested by Wichmann (1999), we estimated the deviance distribution for each model fit by means of 10,000 simulated data-sets for an observer whose correct responses in our experiment are binomially distributed as specified by the model fit. From these simulations, we derived critical values for each reported fit. These values indeed often deviate in an unpredictable manner from the asymptotically derived values, confirming Wichmann (1999). Of course, these critical values do not take into account the non-stationariness of real observers. Overdispersion may thus still occur.

Each deviance residual d_i, is defined as the square root of the deviance value calculated for data point i in isolation, signed according to the direction of the arithmetic residual $y_i - p(\Delta x_i, x_i)$. For binomial data, this is expressed by equation 3-6.

$$d_{ji} = \operatorname{sgn}\left(y_{ji} - p\left(\Delta x_{ji}, x_j\right)\right)\sqrt{2\left[n_{ji}y_{ji}\log\left(\frac{y_{ji}}{p\left(\Delta x_{ji}, x_j\right)}\right) + n_{ji}\left(1-y_{ji}\right)\log\left(\frac{1-y_{ji}}{1-p\left(\Delta x_{ji}, x_j\right)}\right)\right]}. \quad (3\text{-}6)$$

Note that $D = \sum_{j=1}^{Z}\sum_{i=1}^{K_j} d_{ji}^2$, as for RMSE. Systematic trends in deviance residuals indicate a systematic misfit of the model.

3.4.4 Selection

Model selection refers to the problem of selecting, from a group of competing models, the model that best predicts future data, i.e., that generalizes best. Due to noisiness of the data and the problem of over-fitting, this is not simply a matter of goodness-of-fit. In order to select the model with highest *predictive accuracy*, different quantitative methods have been suggested (for a review and overview see Myung, 2000; Pitt & Myung, 2002; Pitt, Myung & Zhang, 2002; Wasserman, 2000; Zucchini, 2000). As there is no generally agreed consensus as to what method is best, we used three different model selection criteria: *Akaike's Information Criterion* (AIC), the *Bayesian information criterion* (BIC) and *cross-validation* (CV). AIC trades simplicity and goodness-of-fit for nested models. It is commonly formulated for model family **F** as given by equation 3-7.

$$\mathrm{AIC}(\mathbf{F}) = D + 2l, \tag{3-7}$$

With l the number of adjustable parameters[2]. As for all model selection methods mentioned, the model that minimizes the criterion should be selected. AIC has the additional advantage that, for nested models, the reduction in AIC can be compared to a χ^2-distribution with degrees-of-freedom equal to the difference in number of free parameters between the models.

For BIC the complexity measure is not only sensitive to the number of adjustable parameters, but it is also modified by sample size (see equation 3-8). Thus BIC is stricter than AIC for $\ln(n) > 2$, i.e., $n > 7$.

$$\mathrm{BIC}(\mathbf{F}) = D + l\ln(n). \tag{3-8}$$

In machine learning the standard model selection criterion is CV (Hastie, Tibshirani & Friedman, 2003; for its use in psychology see e.g. Browne, 2000). We used 10-fold CV: here the data are divided into ten subsamples of equal size. The model is fitted

[2] Note that this is not as straightforward as it may sound, as functions with nominally the same number of free parameters may have inherently more or less complexity. For a non-technical introduction see Forster, 1999; 2000 and Forster & Sober, 1994.

to nine subsamples, the training set. The normalized deviance of the model fit to the training set is called *training error*. The normalized deviance of the same parameter estimates to the subsample that was left out during parameter estimation (i.e., the test set) is called *test error* (see equation 3-9). By minimizing test error, CV has a strong and intuitively appealing emphasis on generalizability (and large differences between training and test error are indicative of over-fitting).

$$\mathrm{CV}(\mathbf{F}) = D_{Test}. \tag{3-9}$$

In the fitting of each training set, lapse rate λ, was re-estimated for each pedestal and noise combination. Each subsample can be used once as test set, which results in ten estimates of test error. Assuming stationary data, i.e., training and test data come from the same distribution, a model that is correct – in particular does not over-fit – has a test error equal to the training error. We did ten iterations of 10-fold CV, leading to 100 parameter estimates and their associated test errors.

3.5 Modelling results I – Simultaneous fit to the data of all noise conditions

To test the null-hypothesis that contrast discrimination is invariant to the presence of noise, once corrected for absolute visibility, each observer's data of both noise conditions were brought to the scale of the noiseless data (i.e., all pedestal and increment contrasts of both noise conditions were normalised by the relevant 75%-correct detection threshold in noise and rescaled by the noiseless detection threshold). This operation transforms both noise detection thresholds to be equal to the noiseless detection threshold. If the invariance-hypothesis holds, this rescaling operation should also remove all systematic differences between conditions. For each observer, the four free parameter gain-control model was fitted to all rescaled data at once. An example of such a fit is shown in Figure 3-2a for observer L.V..

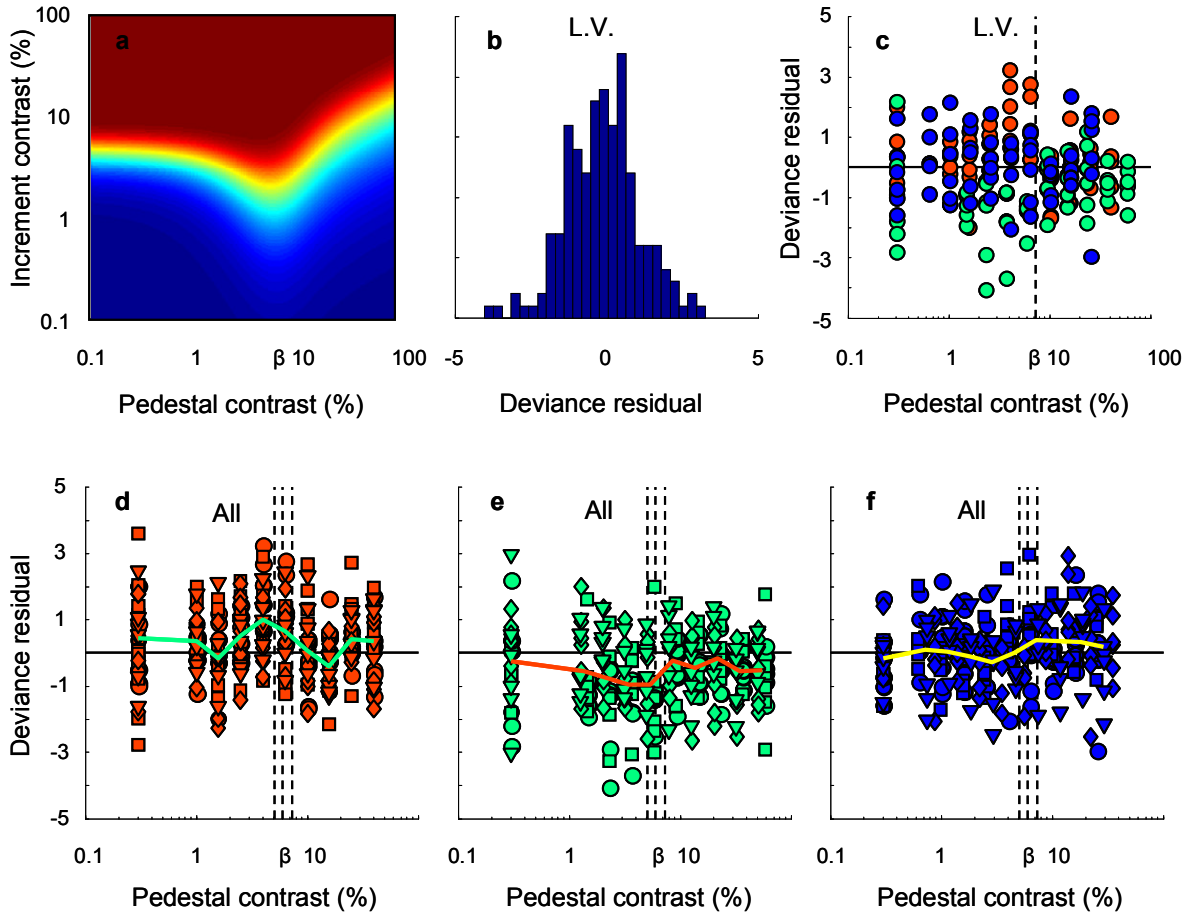

Figure 3-2. ***a.*** *Performance (i.e., percentage correct responses, ranging from 50% –dark blue– to 100% –dark red–) of observer L.V. as a function of pedestal and increment contrast according to the best simultaneous fit of the gain-control model to all noise conditions. As can be seen, the trough of the dip is located at a pedestal contrast equal to the semisaturation contrast β. Note that psychometric functions, i.e., vertical slices, are most shallow in the trough of the dip and steepest for detection.* ***b.*** *The distribution of deviance residuals for observer L.V..* ***c.*** *Deviance residuals as a function of pedestal contrast for observer L.V.. Different colours indicate different noise conditions (see Figure 3-1a). The vertical dashed line depicts β.* ***d.-f.*** *Deviance residuals of the simultaneous fits as a function of pedestal contrast for all observers. Different colours indicate different noise conditions; different symbols indicate different observers (see Figure 3-1a). The vertical dashed lines depict the β-estimates of the different observers. The thick line represents the average deviance residual at each pedestal contrast.*

As can be seen in this figure, the TvC-functions derived from the model fit are dipper-shaped. Furthermore, the dependence of threshold reduction on performance level is captured by the gain-control model. At 75% correct, maximal detection

threshold reduction for observer L.V. is estimated to be 2.7, at 60% correct this number equals 5 and at 90% correct it is 1.9. Figure 3-2b shows the distribution of deviance residuals for this fit, based on 180 blocks of 50 trials each. In Figure 3-2c deviance residuals are plotted as a function of pedestal contrast with different colour labels for different noise conditions.

	α	β	η	κ	$\boldsymbol{\sigma}$	D_{Total}	D_{NN}	D_{WN}	D_{MN}
L.V.	17.95	0.074	2.87	2.39	**1**	1.51**	1.36	2.01**	1.18
E.G.	10.98	0.050	3.70	3.19	**1**	1.52**	1.88**	1.74**	.93
B.B.	16.99	0.059	3.31	2.69	**1**	1.20	1.17	1.39	1.03
L.V.E.	10.34	0.051	3.61	3.14	**1**	1.17	1.13	1.24	1.14

Table 3-1. *Parameter estimates and normalised deviance for the simultaneous fit of the gain-control model to all noise conditions at once. Bold symbols and numbers indicate frozen parameter values. * indicates that D is outside the 99% confidence interval of the deviance of a stationary observer, ** refer to the 99.9% confidence interval.*

Table 3-1 lists the parameter estimates, normalised total deviance and normalised deviance split by noise condition for all observers. Parameter estimates are within the range of typical values. The normalised total deviance values, all ranging between 1.17 and 1.52, indicate a reasonable general quality of fit: the gain-control model thus explains much of the variance in these data. Nevertheless, it should be noted that for two of four observers, total deviance does not belong to the 99.9% confidence interval of the deviance expected if observers were to behave stationary. Calculating deviance by noise condition further reveals that the quality of fit differs over noise conditions: deviance is always lowest for the moderate-noise condition, but may be considerably higher for the no- and weak-noise condition (see Table 3-1).

To explore whether differences between noise conditions underlie this pattern, deviance residuals split by noise condition were analysed by means of linear regression. Deviance residuals of all observers were pooled for this analysis. Figures 3-2(d-f) show the deviance residuals of all observers split by noise level. The thick line

describes the mean deviance residual as a function of pedestal contrast. Judged by eye, the deviance residuals of the no-noise condition (Figure 3-2d) seem to rise from detection threshold till around β – i.e., the semi-saturation contrast if both exponents are equal, corresponding to the location of the trough of the dip, indicated by the vertical dashed lines – while the deviance residuals of the weak-noise condition (Figure 3-2e) display a decreasing trend in the same region. The deviance residuals of the moderate-noise condition (Figure 3-2f) do not display any systematic trend.

Indeed, a linear regression analysis relating the logarithm of pedestal contrast to deviance residual revealed that deviance residuals of the no- and weak-noise condition differ significantly at low pedestal contrasts. Figure 3-3 shows the results of this linear regression analysis. The full lines depict the best fitting linear curves to the data, the dashed lines depict the 99.15% *confidence bands* of these curves and the circles illustrate the mean deviance residual. Deviance residuals of the decreasing and rising part of the dipper function were analysed separately. For both pedestal contrast regions, six comparisons are interesting to make: Do these three lines differ from 0, indicating a systematic misfit of the model? And do they differ from each other, indicating systematic differences between noise conditions? The overall probability of making a Type I error, i.e., falsely rejecting the null-hypothesis, thus equals 0.05 (i.e., $1 - (0.9915)^6$) for both the decreasing and rising part of the dipper function[3].

As can be seen in Figure 3-3, deviance residuals of the no-noise and weak-noise condition differ from zero and from each other for almost the whole decreasing part of the dipper function, while deviance residuals of the moderate noise condition are centred around zero for these pedestal contrasts. These systematic trends indicate that for low pedestal contrasts, contrast discrimination performance in the no-noise condition is underestimated by the model fit to all conditions at once, i.e., the depth of the dip is underestimated. For the weak-noise condition, on the other hand,

[3] Detection data were omitted from the low-pedestal contrast analysis. The reasons are twofold. First, due to the normalisation procedure, most differences between deviance residuals as a function of pedestal contrast will be removed for detection. Second, because of the logarithmic transformation of signal contrast, there is no correct location for these deviance residuals on the contrast-axis.

performance, and consequently the depth of the dip, is overestimated. Making use of a computational model fitted to all data thus confirms the trend suggested by the 75%-correct thresholds: Addition of both weak and moderate noise significantly – in a statistical sense – reduces the size of the pedestal effect. Furthermore, considered across all observers, the reduction seen in the presence of weak noise is more severe than in moderate noise. Modelling our data within the contrast gain-control framework thus allows us to draw statistically sound conclusions which simple confidence-intervals on our raw data failed to unearth.

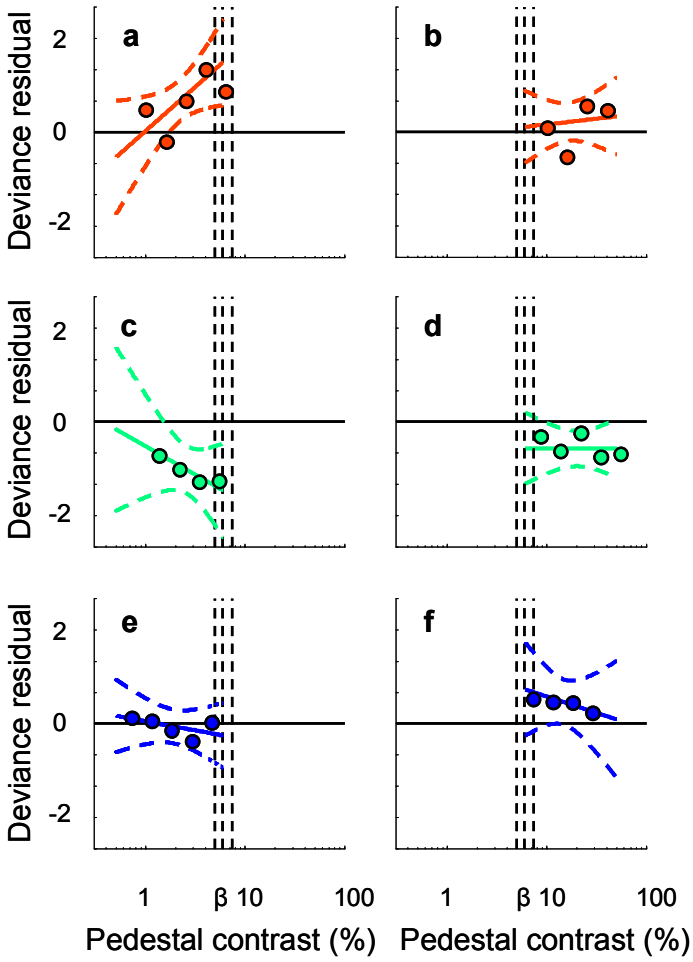

Figure 3-3. *Results of a standard linear regression procedure, relating the logarithm of pedestal contrast to the deviance residuals of the simultaneous fit, split by noise condition. The left figure panels show the results for pedestal contrasts smaller than the semi-saturation contrast β, the right panels for pedestal contrasts larger than β. Full lines indicate the best fitting linear curves; dashed lines indicate the 99.15% confidence band; symbols indicate the average deviance residual.* ***a.-b.*** *The no-noise condition.* ***c.-d.*** *The weak-noise condition.* ***e.-f.*** *The moderate-noise condition.*

At higher pedestal contrasts, deviance residuals of the weak-noise condition are slightly below zero. This may indicate that weak noise affects contrast processing only at low pedestal contrasts. If this were the case, normalising and rescaling the data by means of the detection threshold would slightly misplace the data gathered at higher contrasts. Contrast discrimination performance would be overestimated in this region, as is borne out by our data. No other differences can be noted at higher pedestal contrasts.

In sum, fitting the rescaled data of all noise conditions at once leads to a perhaps reasonable overall quality of fit, but the occurrence of systematic trends in the deviance residuals implies that the null-hypothesis that contrast discrimination is invariant to the presence of noise must be rejected. Though these analyses were run on the pooled deviance residuals, data of all observers displayed the trends described above. To analyse the differences between noise conditions in more detail, we fitted the gain-control model to the data of each noise condition separately.

3.6 Modelling results II – Separate fits to the data of each noise condition

The most parsimonious modification of the gain-control model is to allow the response exponent and gain-control exponent to vary over noise conditions while the response gain and semisaturation contrast are frozen to the estimates of the fit to all data. Indeed, the exponents in the generalised Naka-Rushton equation determine the depth of the dip. Freezing the response gain parameter (a scaling parameter) guarantees that all models operate on the same scale and are thus easily comparable to each other and to the fit to all data at once. Freezing the semi-saturation contrast forces the dip to have the same location for all conditions, i.e., for the 75%-performance contour at a contrast around the 75%-correct detection threshold. In the following, the gain-control model with two effective free parameters, namely the response- and gain-control exponent η and κ, was fitted to each noise condition separately for each observer. An example of these fits can be seen in the upper row of Figure 3-4 for observer E.G.

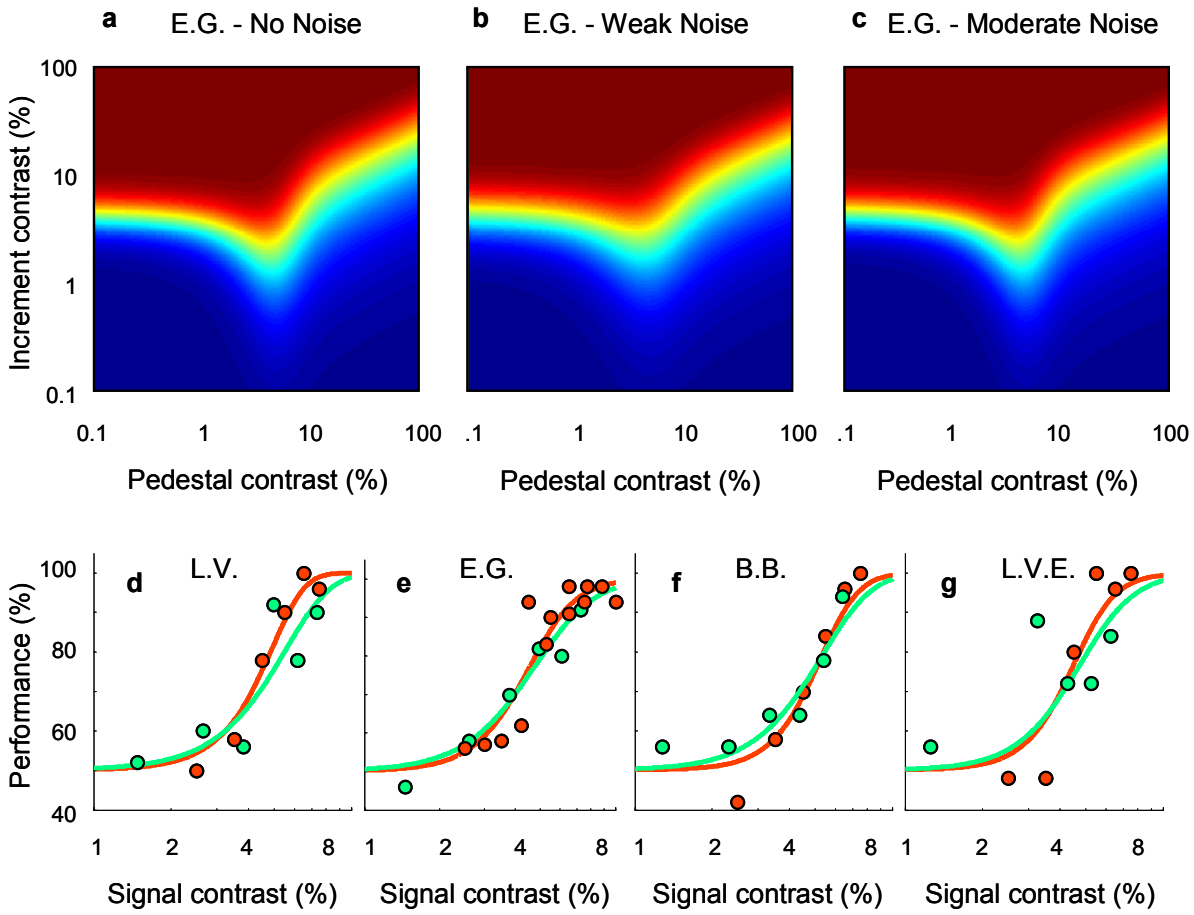

Figure 3-4. ***a.-c.*** *Performance (i.e., percentage correct responses, ranging from 50% – dark blue – to 100% – dark red –) of observer E.G. as a function of pedestal and increment contrast according to the best separate fits of the gain-control model to the different noise conditions.* ***d.-g.*** *The psychometric functions of all observers at zero pedestal contrast for the no-noise and weak-noise condition, according to the best model fits.*

As can be seen from these plots, allowing the exponents to vary over conditions indeed leads to different estimates of the depth of the dip for different noise conditions. It is further noticeable that, in the absence of a pedestal, psychometric functions are estimated to be steeper without noise than in the presence of weak noise for each observer (see Figures 3-4(d-g)), as was also borne out by our data (see Figure 3-1a). The psychometric functions derived from the model fit are plotted in the bottom row of Figure 3-4. Bear in mind that, due to the rescaling operation, psychometric functions are supposed to coincide at 75% correct.

No noise	**α**	**β**	η	κ	**σ**	D
L.V.	**17.95**	**0.074**	3.40	3.00	**1**	1.06
E.G.	**10.98**	**0.050**	3.72	3.24	**1**	1.87**
B.B.	**16.99**	**0.059**	4.17	3.55	**1**	1.00
L.V.E.	**10.34**	**0.051**	4.18	3.75	**1**	1.06
Weak noise						
L.V.	**17.95**	**0.074**	2.79	2.22	**1**	1.13
E.G.	**10.98**	**0.050**	2.99	2.47	**1**	1.44
B.B.	**16.99**	**0.059**	3.06	2.41	**1**	1.37
L.V.E.	**10.34**	**0.051**	3.42	2.92	**1**	1.21
Moderate noise						
L.V.	**17.95**	**0.074**	2.68	2.23	**1**	1.16
E.G.	**10.98**	**0.050**	3.91	3.39	**1**	.93
B.B.	**16.99**	**0.059**	2.94	2.33	**1**	1.00
L.V.E.	**10.34**	**0.051**	3.17	2.70	**1**	1.11

Table 3-2. *Parameter estimates and normalised deviance for the fit of the gain-control model to the separate noise conditions. Bold symbols and numbers indicate frozen parameter values. * indicates that D is outside the 99% confidence interval of the deviance of a stationary observer, ** refer to the 99.9% confidence interval.*

Table 3-2 lists the parameter estimates and normalised deviance for all observers. Comparing the normalised deviance values of Table 3-2 with those of Table 3-1 reveals that quality of fit improved for each noise condition and all observers. It can be seen that normalised deviance is noticeably high for observer E.G. in the no-noise condition (i.e., $D = 1.87$). Fitting this condition with an expanded 6-free parameter version of the gain-control model (i.e., one signal-dependent noise source, having both a multiplicative and exponential component, was added) improved quality of fit only marginally to 1.80, which is not significantly better according to AIC or BIC. Most likely, this data-set is over-dispersed, i.e., observer E.G. displayed non-stationary

behaviour in the no-noise condition (this, of course, cannot be fixed by *any* other model: the "error" is intrinsic to the dataset).

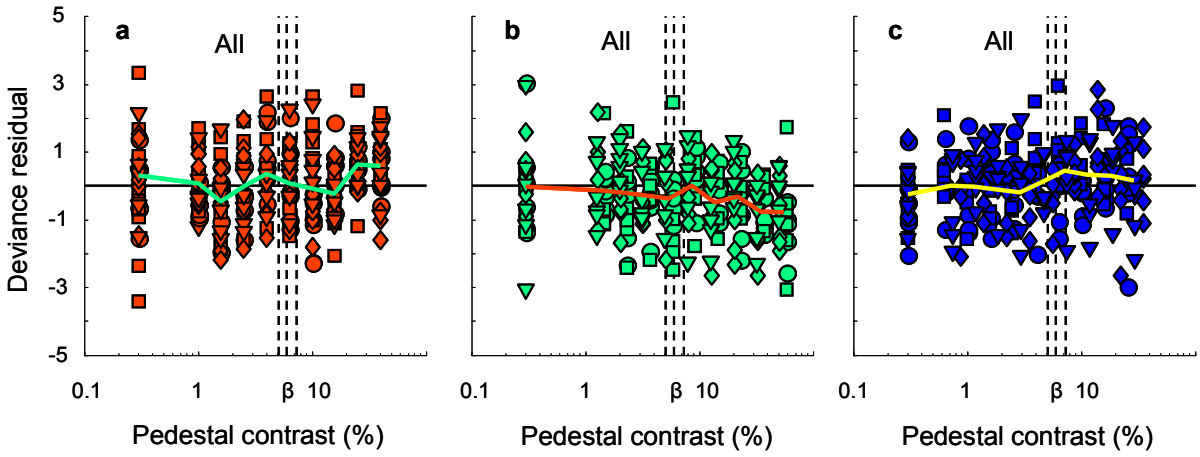

Figure 3-5. *Deviance residuals of the separate fits as a function of pedestal contrast for all observers. Different colours indicate different noise conditions; different symbols indicate different observers. The vertical dashed lines depict the β-estimates of the different observers. The thick line represents the average deviance residual.*

As for the model fit to all conditions at once, we analysed the pooled deviance residuals of all observers by means of a linear regression analysis relating deviance residual to the logarithm of pedestal contrast. The raw data of this analysis are plotted in Figure 3-5, the summary of this analysis in Figure 3-6. For the decreasing part of the dipper, all differences between noise conditions have vanished. As we hypothesized, allowing the response- and gain-control exponent to vary over noise conditions is sufficient to capture the systematic differences between conditions at low pedestal contrasts. At higher contrasts, virtually nothing has changed, so there is still a small but systematic misfit for the weak-noise condition.

In sum, freezing the response gain (α) and semisaturation contrast (β) to the estimates of the fit to the pooled noise conditions and leaving the exponents (η and κ) free to capture the differences between the noise conditions leads to a parsimonious model (i.e., .033 free parameters per block of 50 trials which corresponds to about 1 free parameter per 1,500 trials) that successfully describes our data. Compared to the simultaneous fit to all data, allowing the exponents to vary over noise conditions

leads to an improvement in quality of fit and the disappearance of systematic trends in the deviance residuals at low pedestal contrasts. We now assess whether the improvement brought about by more free parameters is sufficiently large as assessed by methods of model selection.

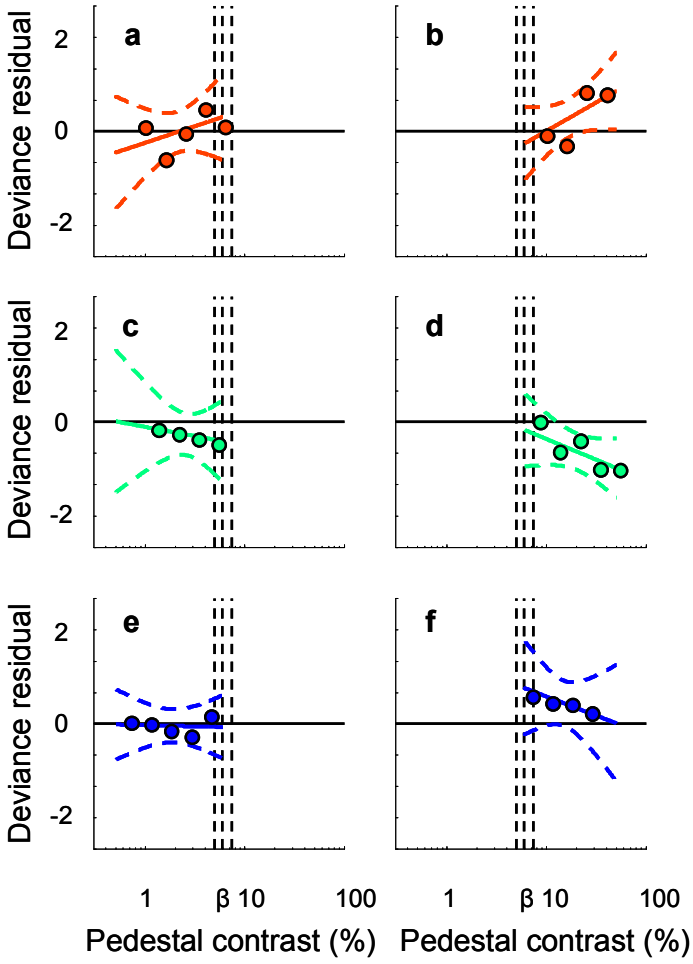

Figure 3-6. *Results of a standard linear regression procedure, relating the logarithm of pedestal contrast to the deviance residuals of the separate fits, split by noise condition. The left figure panels show the results for pedestal contrasts smaller than the semi-saturation contrast β, the right panels for pedestal contrasts larger than β. Full lines indicate the best fitting linear curves; dashed lines indicate the 99.15% confidence band; symbols indicate the average deviance residual.* ***a.-b.*** *The no-noise condition.* ***c.-d.*** *The weak-noise condition.* ***e.-f.*** *The moderate-noise condition.*

3.7 Model selection: simultaneous vs. separate fits

As explained in the paragraph on model selection, there are different ways to assess which of the two modelling approaches has the highest predictive accuracy. We first consider the outcome of the AIC-procedure in detail. First, and most general, we may consider the overall predictive accuracy for all observers and all conditions, analogue to pooling of data across observers and conditions. We must thus compare the AIC of a 16-free parameter model (i.e., 4 x 4 free parameters) to the AIC of a 24-free parameter model (i.e., 4 x 3 x 2 adjustable parameters) for a data-set consisting of 727 blocks. For the fit to all conditions at once, AIC is 992.92; for the separate fits AIC is 893.19. The reduction in AIC thus equals 99.73. This value is notably higher than any sensible critical value derived from a χ^2-distribution with eight free parameters (e.g., 99.9% of the area of this distribution is located below 26.125).

We may thus conclude that, considered across noise conditions and observers, the response gain and semisaturation contrast may be frozen, but it is better not to freeze the exponents of the Naka-Rushton equation. Second, we can also do this analysis for each noise condition, across observers; and for each observer, across noise conditions. And finally, we can do this analysis for each condition within each observer. Results of these analyses are summarised in Table 3-3. Each cell denotes an analysis at the third, most detailed level. Each marginal total denotes an analysis at the second, intermediate level of detail. Significant reductions in AIC are marked by means of stars in Table 3-3.

	No Noise	Weak Noise	Moderate Noise	$\sum$
L.V.	16.78***	51.01***	0.36	68.15***
E.G.	-2.16	17.27**	-0.80	14.31***
B.B.	9.45**	1.14	1.02	11.61**
L.V.E.	3.37	1.25	1.04	5.66

Table 3-3. *Reduction in AIC for each observer and each noise condition. * indicates a significant reduction at $\alpha = 0.05$, ** refers to $\alpha = 0.005$ and *** to $\alpha = 0.001$. Whenever the theoretical degrees-of-freedom were not a natural number (e.g., 2.66), we used the conservative critical value of the χ^2-distribution of the next natural number (i.e., 3).*

It will be noted that the second model is favoured over the first for three of four observers. For observer L.V.E., there is a trend in the same direction, but the reduction in AIC is only marginally significant ($AIC_1 - AIC_2 = 5.66$; $p < 0.059$). We suspect this may at least in part result from the relative lack of data which reduces statistical power – she completed only 4,500 trials – because her deviance residuals and parameter estimates are not inconsistent with other observers. When considering the different noise conditions, it is clear that the improvement in predictive accuracy is mainly due to the better fits to the weak-noise condition ($AIC_1 - AIC_2 = 70.67$; $p < 10^{-5}$) and the no-noise condition ($AIC_1 - AIC_2 = 27.44$; $p < 10^{-5}$). For the moderate-noise condition, fits were already fine in the first approach, so not much could be gained ($AIC_1 - AIC_2 = 1.62$; $p < .65$).

	No Noise	Weak Noise	Moderate Noise	∑
L.V.	15.50	49.73	-0.91	64.32
E.G.	-3.44	16.00	-2.07	10.49
B.B.	8.18	-0.13	-0.25	7.8
L.V.E.	2.56	0.44	0.23	3.23
∑	22.81	66.04	-3	

Table 3-4. *Reduction in BIC for each observer and each noise condition.*

Table 3-4 summarizes the reduction in BIC. Positive numbers indicate that the separate fits should be selected, negative numbers that the simultaneous fit should be selected. As can be seen in Table 3-4, conclusions of the BIC-analysis mostly agree with the AIC-analysis. At the most general level, i.e., considering the pooled data of all observers and noise conditions, the second modelling approach does much better than the first, according to BIC. At the level of individual observers, the same trend is obvious. At the level of noise conditions, the no-noise and weak-noise condition clearly benefit from the additional free parameters in the separate fits. This is not the case for the moderate noise condition, for which BIC selects the simultaneous fit.

	Model I	Model II	Model I-Model II
L.V.	1.54**	1.16	0.39
E.G.	1.54**	1.47**	0.07
B.B.	1.24*	1.15	0.09
L.V.E.	1.19	1.16	0.04
∑	1.38	1.24	0.15

Table 3-5. *Average test error for each observer across noise conditions for the simultaneous and separate fits and their difference. * indicates that the average D_{Test} is outside the 99% confidence interval of the deviance of a stationary observer, ** refers to the 99.9% confidence interval. For Model II, the presence of stars indicates that the average D_{Test} was outside the relevant confidence interval for at least one noise condition.*

Table 3-5 summarizes the average CV-index (i.e., *test error*) and the difference in test error for both modelling approaches for each observer. As indicated by the positive differences in test error, the separate fits provide better predictions for *unseen* data for each observer. It should in addition be noted that for three of four observers, the average test error of each noise condition belongs to the 99% confidence interval of the distribution of deviance values of a stationary observer: the data are thus very well described by our model.

In sum, despite differences in the aspects of model complexity captured by AIC, BIC and CV, all support the same conclusion: When pooled data and observers are considered, the separate fits are always selected over the joint fit. This conclusion also holds at the level of individual observers. At the level of the noise conditions, the improved predictive accuracy is a result of the better fit to both the no-noise and weak-noise condition.

3.8 Discussion

Given our comprehensive model selection we are now in a position to use the parameter estimates of the cross-validation to assess differences between noise conditions. Because the response gain and semisaturation contrast were frozen, we only need to consider the exponents of the Naka-Rushton function. Box plots of the estimates of the response exponent η, the gain-control exponent κ and their difference are shown in Figure 3-7.

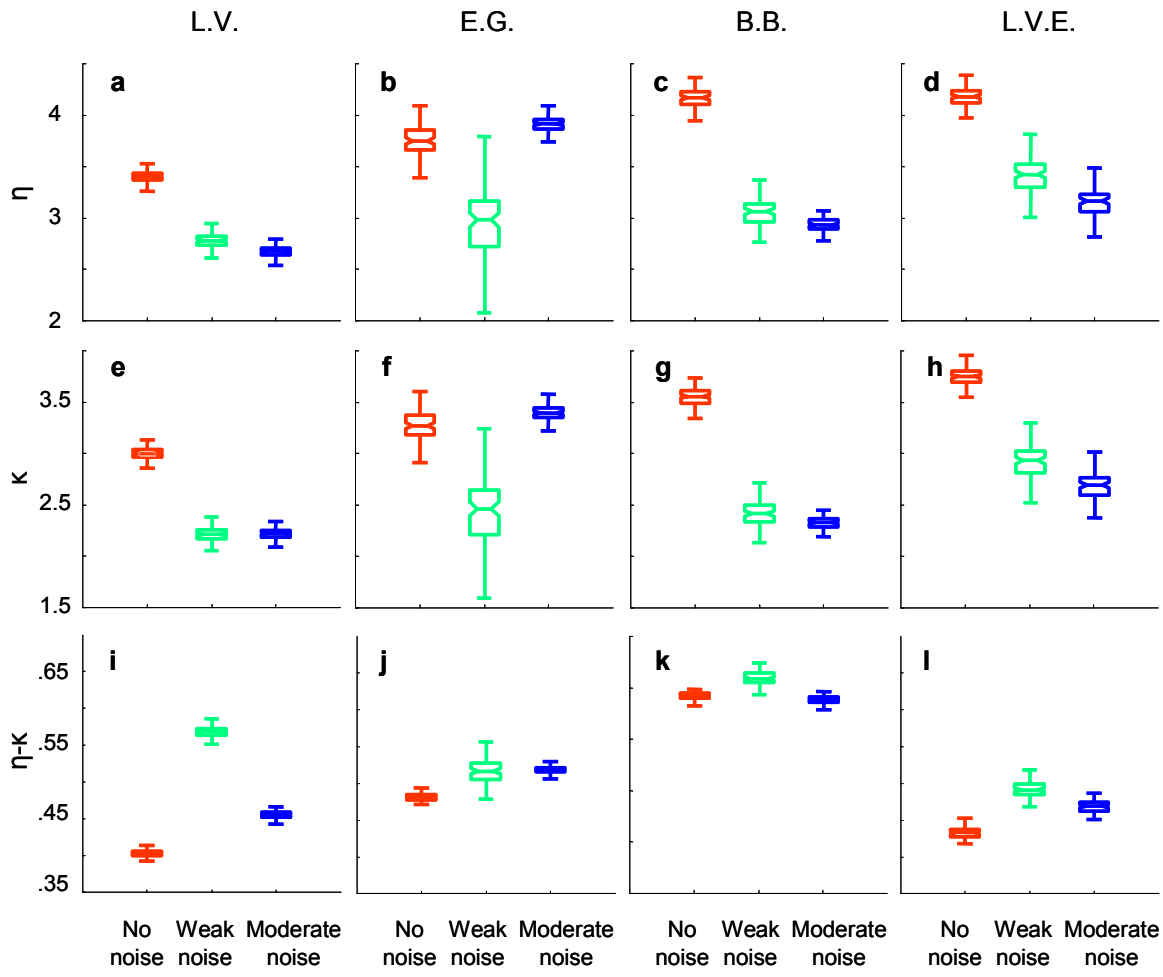

Figure 3-7. ***a.-d.*** *Box plots of the 100 η-values, i.e., the response exponent, estimated in the cross-validation for all observers.* ***e.-h.*** *Box plots of the 100 κ-values, i.e., the gain-control exponent, estimated in the cross-validation for all observers.* ***i.-l.*** *Box plots of the 100 differences between η and κ estimated in the cross-validation analysis for all observers. In all these box plots, the central horizontal line indicates the second quartile (i.e., the median), while the other horizontal lines indicate the first and third quartile. Notches around the median indicate the 95% confidence interval of the median. Whiskers indicate one and a half times the interquartile range. Outliers have been omitted for clarity.*

We first compare parameter estimates for the no-noise and weak-noise condition. For all observers, both exponents are estimated to be reduced in the presence of weak noise. Furthermore, the difference between the exponents is always higher for the weak-noise condition than for the no-noise condition. It is this difference, together with the absolute value of the response exponent, which determines the strength of the pedestal effect: The smaller the difference and the larger the response exponent, the bigger the pedestal effect is. In other words, the depth of the dipper function is reduced in the presence of weak noise for all observers.

It is interesting to note that results are not as systematic for the moderate-noise condition. For some observers, exponents are estimated to be reduced relative to the no-noise condition (e.g., observer B.B.), but for others, this is clearly not the case (e.g., observer E.G.). A similar variability over observers is present in the differences between both exponents: for some this difference has increased in the presence of moderate noise (e.g., observer L.V.E.), but for others this is not the case (e.g., observer B.B.). This variation over observers is not inconsistent with data-sets that have been published earlier: the dipper function of some observers seems to be invariant to the presence of strong noise, while this is not the case for others (e.g., Henning & Wichmann, 2007).

We thus find three stable differences between the no noise and weak-noise condition, namely a reduction in both the response- and gain-control exponent and an increased difference between these exponents in the presence of weak noise. This indicates that the pedestal effect was reduced for all observers in the presence of weak noise. For the moderate-noise condition, results vary over observers.

To better understand these differences between noise conditions, it is helpful to 'open' the models and have a look at the internal contrast response functions (see Figure 3-8; cf. Kienzle, Wichmann, Schölkopf & Franz, 2007; Wichmann, Graf, Simoncelli, Bülthoff & Schölkopf, 2005). Plotted on double logarithmic coordinates, internal response rises steeply until β, the semi-saturation contrast. In this region, the contrast response behaves as an accelerating non-linearity. Because internal variance

is constant at all contrast levels in our fits ($\sigma = 1$), these functions could also be interpreted as detection functions (signal-to-noise ratio as a function of contrast). Indeed, detection sensitivity has been reported to rise in an accelerating way as a function of contrast (Foley & Legge, 1981; Nachmias, 1981; Nachmias & Sansbury, 1974). It is this acceleration that leads to the response expansion that underlies the pedestal effect. The larger the log-log steepness at low contrasts, the stronger the pedestal effect.

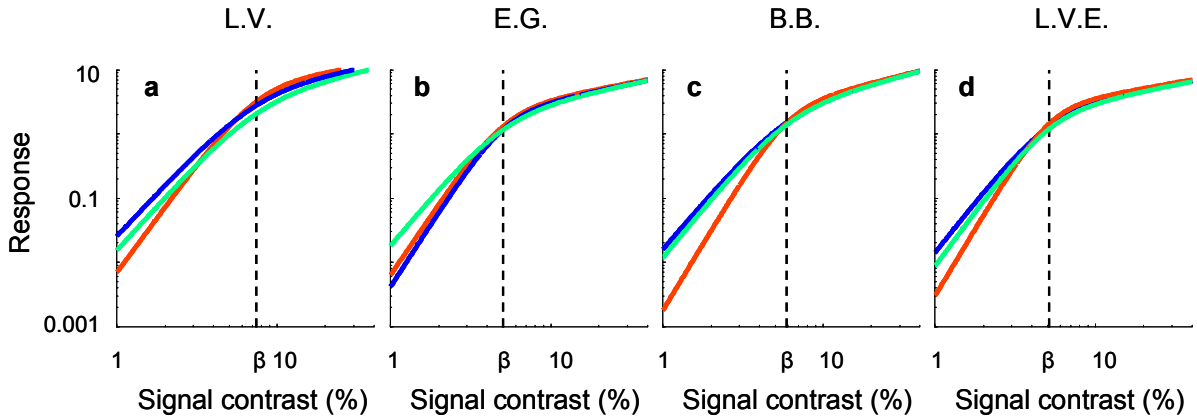

Figure 3-8. *Contrast response functions according to the best separate fits of the gain-control model to the different noise conditions, plotted on double logarithmic coordinates. Red lines refer to the no-noise condition, green to the weak-noise condition and blue to the moderate-noise condition.*

Comparing the no-noise condition to the weak-noise condition illustrates that the log-log steepness, and thus the pedestal effect, is reduced for all observers in the presence of weak noise. The reduced log-log steepness in weak noise is a consequence of the higher response at low contrasts, which in turn leads to improved sensitivity. The response difference between the no-noise and weak-noise condition diminishes as a function of contrast and disappears completely around the semi-saturation contrast (due to the rescaling procedure).

This is not inconsistent with the effect of a rectification mechanism at the output of a linear filter stage, prior to the non-linear response expansion. Due to rectification, the mean response of a linear filter is enhanced in the presence of noise. At zero pedestal

contrast, the response-enhancing effect of rectification is maximal, because half of the responses of a non-rectified linear filter are negative. As pedestal contrast, and thus the average filter response, increases, the proportion of negative responses drops and the response-enhancing effect of rectification diminishes until it vanishes completely. This is consistent with the contrast response functions plotted in Figure 3-8.

Thus, using the gain-control model to get a small parameter description of our data that is statistically sound indeed shows that some parameters change in the presence of noise. We do not wish to claim that these parameters actually change in the visual system, but rather that this is the statistical signature of how noise changes the data.

To further explore whether one mechanism may underlie the data, we simulated performance of the Goris et al. gain-control model in the contrast discrimination experiments reported here (see Figure 3-9, upper row). In this simulation, the filter stage – which need not be specified explicitly for sinusoidal contrast discrimination – was chosen to consist of *(optimal) template matching*, followed by full-wave rectification (e.g., Lu & Dosher, 2008). Template matching is a convenient way to transform 2-D input images to 1-D 'filter responses'. Prior to this filter stage, *stimulus sampling* or limited calculation efficiency was assumed, as is often the case in detection-in-noise models (e.g., Lu & Dosher, 2008). This is described by parameter k, which expresses the proportion of available information used by the observer and ranges between 0 and 1. Inefficiencies in the visual system need not be conceptualised as sampling: parameter k could be thought of as, e.g., reflecting the use of a suboptimal filter, for instance a spatial-frequency tuned channel that has an effective bandwidth that is broader than the narrowband Gabor signal.

To describe the nonlinear mapping of stimulus contrast to internal contrast representation, the second part of the transduction mechanism consisted of the generalised four parameter Naka-Rushton function (free parameters α, β, η and κ). The rectified filter responses used in the expansive, i.e., the numerator, and the compressive, i.e., the denominator, parts of the Naka-Rushton function were the same.

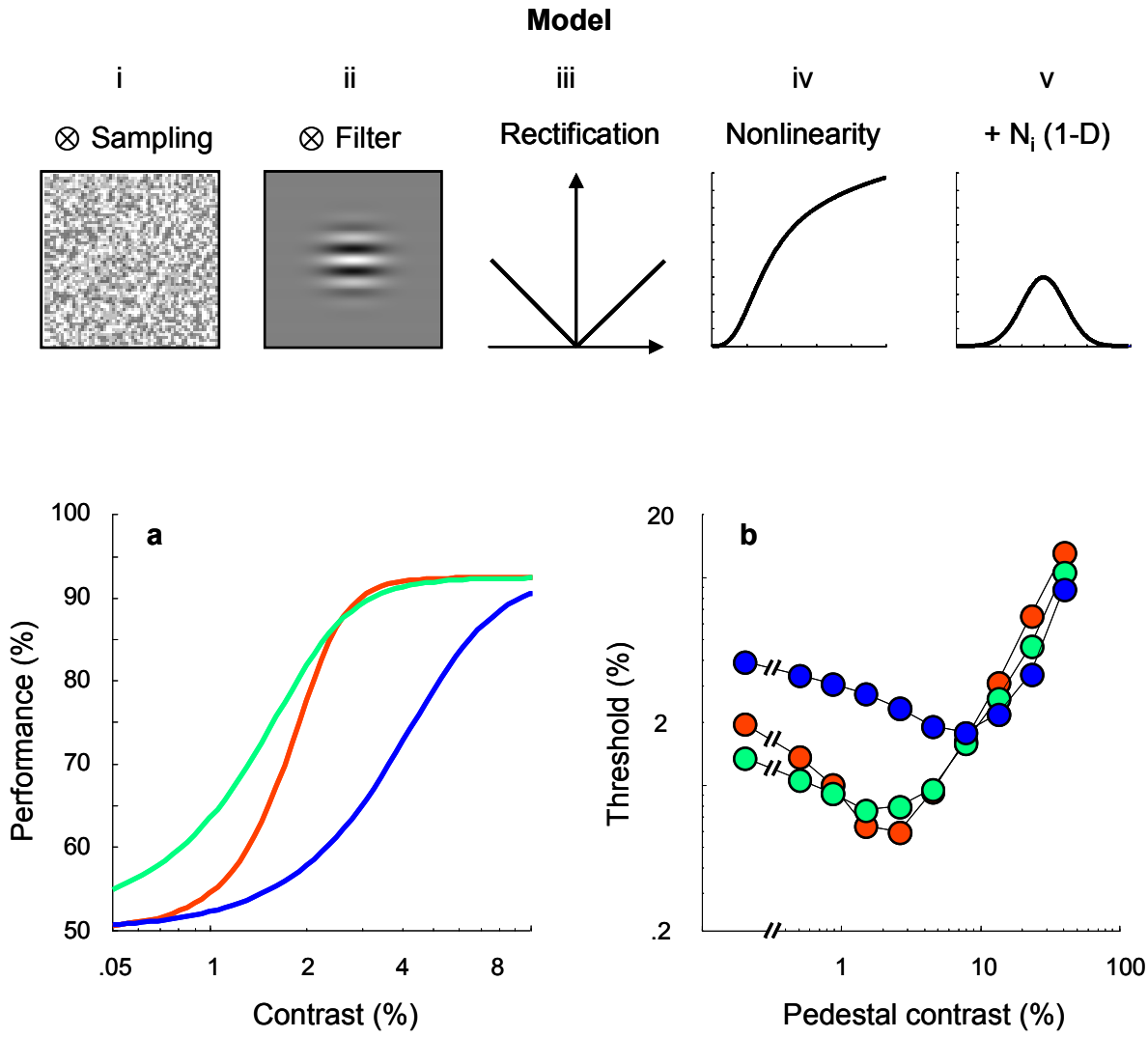

Figure 3-9. ***Upper row.*** *Illustration of the model components used in the simulation. Stimuli with a narrow-band spatial-frequency spectrum are often assumed to be processed in a single spatially localised, spatial-frequency and orientation-selective filter. Ideally, an observer uses a filter that is an exact template of the signal (ii). Human deviations from ideal observer performance are typically interpreted as stemming from internal noise sources (v) on the one hand and limited efficiency or sampling (i) on the other hand (this refers to using only a sample of the available image information, for instance only the image pixels corresponding to the white pixels in i). The transduction mechanism further consisted of rectification (iii) and a nonlinear post-filter stage –the generalised Naka-Rushton transformation– (iv). The final model component was a late, signal-independent noise source (v).* ***Lower row. a.*** *Simulated performance as a function of signal contrast on semi-logarithmic coordinates for the three detection conditions. Red lines refer to the no-noise condition, green to the weak-noise condition and blue to the moderate-noise condition.* ***b.*** *The simulated TvC-functions at 75% correct.*

Although some evidence points to the existence of a broadly tuned contrast gain control pool (e.g., Foley, 1994; Holmes & Meese, 2004), we opted to use only within-channel suppression in this model to avoid an increase of the number of parameters. Furthermore, because the spectral properties of the noise were constant for the weak and moderate noise conditions, it is unlikely that this simplification has a significant impact on our conclusions. The transduction mechanism in this simulation is thus fully determined by specifying the sampling (k) and the parameters of the generalised Naka-Rushton equation (α, β, η and κ).

The variability of internal states is determined by internal noise on the one hand and the use of stochastic stimuli on the other hand. One parameter was used to describe the late 1-D internal noise, assumed to be Gaussian and signal-independent (parameter σ). The effect of the external noise on internal variability was derived from Monte Carlo simulations. With the noise levels used in our experiments as input to the gain-control model, k, η, κ and β were varied – effects of α, i.e. the response gain, need not be simulated –. Descriptive functions were fitted to the simulated response variances. Assuming equal variance of the signal-plus-noise and noise representations (as a first approximation), allowed us to use these descriptive functions to formalise the full model behaviour.

In sum, the model used in this simulation consisted of *image sampling* (k), *template matching*, *response rectification*, *nonlinear transduction* (α, β, η and κ) and *late noise addition* (σ). All these components are illustrated in the upper row of Figure 3-9. As we did in the rest of the paper, σ was frozen to 1. The other five model parameters were chosen in such a way that the (normalised) simulated thresholds would approximate the data shown in Figure 3-1. We found this to be the case for $k = 0.05$ (i.e., sampling), $\alpha = 2*10^6$ (i.e., the response gain), $\beta = 0.025$ (i.e., the semisaturation contrast), $\eta = 3$ (i.e., the response exponent), $\kappa = 1.75$ (i.e., the gain-control exponent) and $\sigma = 1$ (i.e., the late noise).

Figure 3-9a shows the psychometric functions relating detection performance to signal contrast without noise (red) and in the presence of weak (green) and moderate noise

(blue) on semi-logarithmic coordinates. As was the case for the detection in noise data discussed in this paper and shown in Figure 3-1a, addition of weak noise improves contrast detection performance. This is not the case for moderate noise. In weak noise, the 75%-correct detection threshold is reduced by a factor of 1.39 (as compared to 1.39 across observers). In modest noise, this threshold is increased by a factor of 2.14 (1.34 across observers). Further, psychometric functions are not parallel on semi-logarithmic coordinates. With these parameter values, the Goris et al. gain-control model thus mimics some aspects of human detection in noise data.

Figure 3-9b shows the full simulated TvC-functions at 75% correct for all noise conditions on double logarithmic coordinates (compare to Figures 3-1(b-c)). The most leftward points denote detection (i.e., the pedestal contrast equals zero). For these TvC-functions, maximal threshold reduction at 75% correct was of a factor 3.38 without noise (3.44 across observers), 1.78 in the presence of weak noise (2.17 across observers) and 2.23 in the presence of moderate noise (2.62 across observers). As was the case for the discrimination data discussed in this paper and shown in Figures 3-1(b-c), the depth of the dip is thus most reduced in the presence of weak noise. It may further be noticed that contrast discrimination thresholds of the three conditions almost coincide at high pedestal contrasts, consistent with our data. At these contrasts, the slope of the dipper handle approximates one on double logarithmic coordinates, in line with several psychophysical observations (e.g., Bird, Henning & Wichmann, 2002). In summary, a single contrast processing mechanism can produce contrast discrimination in noise data that resemble many aspects of our experimental findings.

Finally, as we did for our observers, the gain-control model used throughout the paper was also fitted to the normalised and rescaled simulated data, split per noise level (in line with the previous model fitting, the response gain and semisaturation contrast were frozen to the estimates of the fit to all data, i.e., $\alpha = 12.3$ and $\beta = 0.04$). Note that this gain-control model is similar, but not identical to the model used to produce these data. The most interesting parameter estimates are shown in Figure 3-10 (compare to Figure 3-7).

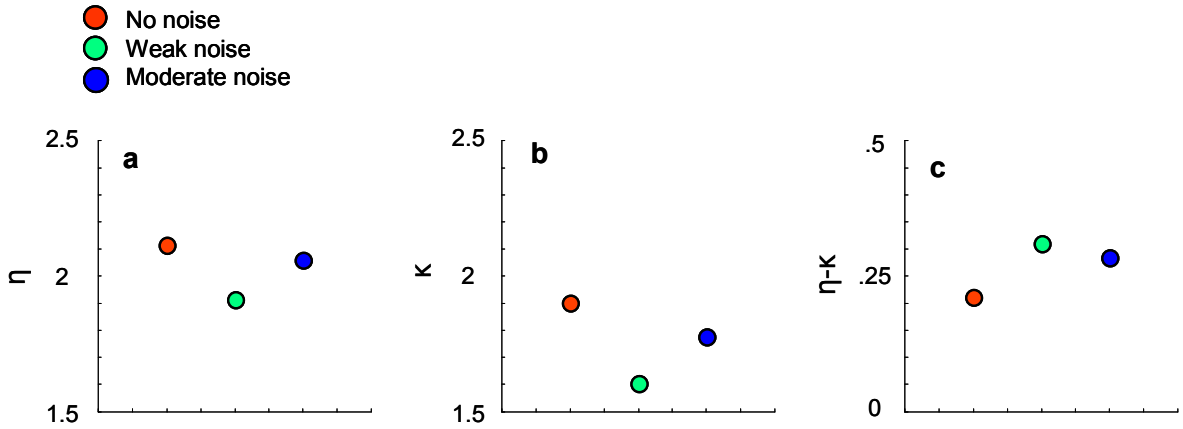

Figure 3-10. ***a.*** *The response exponent estimates for the simulated data set.* ***b.*** *The gain-control exponent estimates for the simulated data.* ***c.*** *The differences between η and κ for the simulated data set.*

It will be noted that the three parameter changes discussed above are also present in the parameter estimates for the simulated data set (i.e., a reduction in both the response- and gain-control exponent and an increased difference between these exponents in the presence of weak noise, relative to no-noise). This simulation thus shows that a single mechanism underlying the pedestal effect and stochastic resonance may have the signature of reduced exponents if the gain-control model is fitted to the data. These similarities further support the interpretation of the contrast discrimination data discussed in this paper as consistent with the idea that a single mechanism underlies the pedestal effect in contrast discrimination and stochastic resonance in contrast detection.

It is harder to explain how contrast processing changes in the presence of moderate noise. A main part of the problem is that data differ considerably between observers. For some observers, the pedestal effect is much reduced in moderate noise. But for others, hardly anything seems to differ. An easy, common interpretation is thus unlikely to be found.

3.9 Conclusion

In this paper we explored whether a single mechanism underlies the pedestal effect, i.e., the improved detectability of a grating in the presence of a low-contrast masking grating, and stochastic resonance, i.e., the improved detectability of a grating in the presence of sub-threshold noise. Analysis of the 75%-correct thresholds was not conclusive, i.e., the trends present in the data were not significant. Making use of a full quantitative description of our data with few parameters, i.e., using the gain-control model as a statistical tool, combined with comprehensive model selection assessments we showed the pedestal effect to be *reduced* in the presence of weak noise for all observers. This reduction clearly rules out independent, additive sources of performance improvement and cannot simply be attributed to additionally introduced response variability by the weak noise, because it was smaller and not as consistent in the presence of moderate noise. We further showed that a single mechanism responsible for the pedestal effect and stochastic resonance may have the signature of reduced exponents if the gain-control model is fitted to the data. Given that the pattern of parameter changes for real data is the same as for simulated data (under the hypothesis of a single mechanism) and that the alternative hypothesis can be ruled out by model selection, we interpret these data as indicating that a single mechanism underlies the pedestal effect and stochastic resonance in contrast perception.

3.10 References

Bach, M. (1997). A note on luminance calibration of raster-scan cathode-ray tubes: temporal resolution, ripple, and accuracy. *Spatial Vision, 10(4)*, 485–489.

Bach, M., Meigen, T., and Strasburger, H. (1997). Raster-scan cathode-ray tubes for vision research limits of resolution in space, time and intensity, and some solutions. *Spatial Vision, 10(4)*, 403–414.

Baker, D. H., Meese, T. S., & Georgeson, M. A. (2007). Binocular interaction: Contrast matching and contrast discrimination are predicted by the same model. *Spatial Vision 20(5),* 397-413.

Bird, C. M., Henning, G. B., & Wichmann, F. A. (2002). Contrast discrimination with sinusoidal gratings of different spatial frequency. *Journal of the Optical Society of America A, 19*, 1267-1273.

Blackwell, K. T. (1998). The effect of white and filtered noise on contrast detection thresholds. *Vision Research, 38(2)*, 267-280.

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

Bradley, A., & Ohzawa, I. (1986). A comparison of contrast detection and discrimination. *Vision Research*, *26*, 991-997.

Brainard, D. H. (1989). Calibration of a computer controlled colour monitor. *COLOR research and application, 14(1)*, 23–34.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433-436.

Browne, M. W. (2000). Cross-validation methods. *Journal of Mathematical Psychology*, *44*,108–132.

Campbell F. W., Carpenter, R. H. S., & Levinson, J. Z. (1969). Visibility of aperiodic patterns compared with that of sinusoidal gratings. *Journal of Physiology (London), 204*, 283-298.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology (London), 197*, 551-556.

Collins, J. J., Imhoff, T. T., & Grigg, P. (1996). Noise-enhanced tactile sensation. *Nature, 383*, 770.

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America. A, 2(7)*, 1160–1169.

Derrington, A. M., & Henning, G. B. (1989). Some observations on the masking effect of two-dimensional stimuli. *Vision Research, 29(2)*, 241-246.

DeValois, R. L., & DeValois, K. K. (1988). *Spatial Vision.* Oxford: Oxford University Press.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: masking experiments require a new model. *Journal of the Optical Society of America A, 11(6)*, 1710–1719.

Foley, J. M., & Chen C.C. (1997). Analysis of the effect of pattern adaptation on pattern pedestal effect: A two-process model. *Vision Research, 37*, 2779-2788.

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

Foley, J. M., & Schwarz, W. (1998). Spatial attention: Effect of position uncertainty and number of distracter patterns on the threshold-versus-contrast function for contrast discrimination. *Journal of the Optical Society of America A, 15*, 1036-1047.

Forster, M. R. (1999). Model selection in science: The problem of language variance. *British Journal for the Philosophy of Science, 50(1)*, 83–102.

Forster, M. R. (2000). Key concepts in model selection: performance and generalizability. *Journal of Mathematical Psychology, 44*, 205–231.

Forster, M. R., & Sober, E. (1994). How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. *British Journal for the Philosophy of Science, 45*, 1–35.

Georgeson, M. A., & Georgeson, J. M. (1987). Facilitation and masking of briefly presented gratings: Time-course and contrast dependence. *Vision Research, 27(3)*, 369–379.

Georgeson, M. A., & Meese, T. S. (2006) Fixed or variable noise in contrast discrimination? The jury's still out… *Vision Research, 46*, 4294-4303.

Gorea, A., & Sagi, D. (2001). Disentangling signal from noise in visual contrast discrimination. *Nature Neuroscience, 4(11)*, 1146-1150.

Goris, R. L. T., Zaenen, P., & Wagemans, J. (2008). Some observations on contrast detection in noise. *Journal of Vision, 8(9): 4*, 1-15.

Graham, N. (1989). *Visual Pattern Analyzers*. Oxford: Oxford University Press.

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: a comparison of single channel and multichannel models. *Vision Research, 11*, 251-259.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics.* New York: John Wiley & Sons, Inc.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2003). *The Elements of Statistical Learning Theory*. Springer.

Henning, G. B. (1988). Spatial-frequency tuning as a function of temporal frequency and stimulus motion. *Journal of the Optical Society of America A, 5*, 1362-1373.

Henning, G. B., Bird, C. M., & Wichmann, F. A. (2002). Contrast discrimination with pulse trains in pink noise, *Journal of the Optical Society of America A, 19*, 1259-1266.

Henning, G. B., Derrington, A. M., & Madden, B. C. (1983). Detectability of several ideal spatial patterns. *Journal of the Optical Society of America, 73*, 851-854.

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

Henning, G. B., Hertz, B. G., & Hinton, J. L. (1981). Effects of different hypothetical detection mechanisms on the shape of spatial-frequency filters inferred from masking experiments: I. noise masks. *Journal of the Optical Society of America, 71*, 574-581.

Henning, G. B., & Wichmann, F. A. (2007). Some observations on the pedestal effect. *Journal of Vision, 7(1:3)*, 1-15.

Holmes, D. J. & Meese, T. S. (2004). Grating and plaid masks indicate linear summation in a contrast gain pool. *Journal of Vision*, *4(12:7)*, 1080-1089.

Kienzle, W., Wichmann, F. A., Schölkopf, B., & Franz, M. O. (2007). A nonparametric approach to bottom-up visual saliency. In Schölkopf, B., Platt, J., and Hoffman, T., editors, *Advances in Neural Information Processing Systems 19*, pages 689–696. MIT Press, Cambridge, MA.

Kontsevich, L. L., Chen, C.-C., & Tyler, C. W. (2002). Separating the effects of response nonlinearity and internal noise psychophysically. *Vision Research, 42*, 1771-1784.

Legge, G. E. (1981). A power law for contrast discrimination. *Vision Research, 21(4)*, 457-467.

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

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

Lu, Z.-L., & Dosher, B. A. (2008). Characterizing observers using external noise and observer models: assessing internal representations with external noise. *Psychological review, 115(1),* 44-82.

Myung, I.-J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psychology, 44*, 190–204.

Nachmias, J. (1981). On the psychometric function for contrast detection. *Vision Research 21,* 215-223.

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

Naiman, A. C., & Makous, W. L. (1992). Spatial non-linearities of grayscale CRT pixels. In *Human Vision, Visual Processing, and Digital Display III, SPIE Proc. 1666.*, pages 41–56. SPIE, Bellingham, WA.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America. A, 2(9)*, 1508-1532.

Pelli, D. G. (1997a) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*, 437-42.

Pelli, D. G. (1997b). Pixel independence: measuring spatial interactions on a CRT display. *Spatial Vision, 10(4)*, 443–446.

Pitt, M. A., & Myung, I.-J. (2002). When a good fit can be bad. *Trends in Cognitive Sciences*, *6*, 421–425.

Pitt, M. A., Myung, I.-J., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review, 109(3)*, 472-491.

Smithson, H. E., Henning, G. B., MacLeod, D. I. A. & Stockman, A. (under review). The effect of notched noise on flicker detection and discrimination.

Wasserman, L. (2000). Bayesian model selection and model averaging. *Journal of Mathematical Psychology, 44*, 92–107.

Watson, A. B., Barlow, H. B., & Robson, J. G. (1983). What does the eye see best? *Nature 302,* 419-422.

Wichmann, F. A. (1999). *Some aspects of modelling human spatial vision: Contrast discrimination.* Unpublished doctoral dissertation, The University of Oxford, Oxford, UK.

Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness-of-fit. *Perception & Psychophysics, 63*, 1293-1313.

Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics, 63*, 1314-1329.

Wichmann, F. A., Graf, A. B. A., Simoncelli, E. P., Bülthoff, H. H., & Schölkopf, B. (2005). Machine learning applied to perception: decision-images for gender classification. In Saul, L. K., Weiss, Y., and Bottou, L., editors, *Advances in Neural Information Processing Systems 17*, pages 1489–1496. MIT Press.

Wiesenfeld, K., & Moss, F. (1995). Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDS. *Nature, 373*, 33-36.

Yang, J., & Makous, J. (1995). Modeling pedestal experiments with amplitude instead of contrast. *Vision Research, 35(14)*, 1979-1989.

Yu, C., Klein, S. A., & Levi, D. M. (2002). Facilitation of contrast detection by cross-oriented surround stimuli and its psychophysical mechanisms. *Journal of Vision, 2*, 243-255.

Zeng, F.-G., Fu, Q.-J., & Morse, R. P. (2000). Human hearing enhanced by noise. *Brain Research, 869,* 251-255.

Zucchini, W. (2000). An introduction to model selection. *Journal of Mathematical Psychology*, *44*, 41–61.

Chapter 4 A neurophysiologically plausible population-code model for human contrast discrimination

The pedestal effect is the improvement in the detectability of a sinusoidal grating in the presence of another grating of the same orientation, spatial frequency and phase – usually called the pedestal. Recent evidence has demonstrated that the pedestal effect is differently modified by spectrally flat and notch-filtered noise: the pedestal effect is reduced in flat noise, but virtually disappears in the presence of notched noise (Henning & Wichmann, 2007). Here we consider a network consisting of units whose contrast response functions resemble those of the cortical cells believed to underlie human pattern vision and demonstrate that, when the outputs of multiple units are combined by simple weighted summation – a heuristic decision rule that resembles optimal information combination and produces a contrast-dependent weighting profile – the network produces contrast-discrimination data consistent with psychophysical observations: the pedestal effect is present without noise, reduced in broadband noise, but almost disappears in notched noise. These findings follow naturally from the normalization model of simple cells in primary visual cortex, followed by response-based pooling, and suggest that in processing even low-contrast sinusoidal gratings, the visual system may combine information across neurons tuned to different spatial frequencies and orientations.

This chapter has been submitted for publication as:

Goris, R. L. T., Wichmann, F. A., & Henning, G. B. (2009). A neurophysiologically plausible population-code model for human contrast discrimination. *Journal of Vision*.

4.1 Introduction

Behavioural evidence from detection experiments that involve barely visible low-contrast stimuli suggests that information in the early visual system is analysed in spatially localised, spatial-frequency-tuned, orientation-selective ´channels` of limited bandwidth (Blakemore & Campbell, 1969; Campbell & Robson, 1968; DeValois & DeValois, 1988; Graham, 1989; Graham & Nachmias, 1971) (see Figure 4-1a). At the neural level, the multidimensional stimulus selectivity of cortical neurons makes primary visual cortex a likely substrate for these channels (DeValois & DeValois, 1988; Graham, 1989).

Understanding the processing of stimuli that are more than just detectable, however, is a prerequisite for any useful model of spatial vision. To gain insight into visual processing of supra-threshold contrasts, sinusoidal contrast discrimination has been studied extensively (Bird, Henning & Wichmann, 2002; Foley, 1994; Goris, Wagemans & Wichmann, 2008; Henning & Wichmann, 2007; Legge & Foley, 1980; Nachmias & Sansbury, 1974; Wichmann, 1999). One of the main findings of contrast-discrimination studies is the pedestal effect – the improved detectability of a sinusoidal "signal" grating in the presence of a low-contrast grating with the same spatial frequency, orientation and phase as the signal and often called "the pedestal".

Characteristics of contrast transduction and/or gain-control mechanisms believed to operate within single channels have often been inferred from the pedestal effect on the assumption that stimuli with narrowband spatial-frequency spectra are processed within a single channel tuned to the orientation and spatial frequency of the signal (Foley, 1994; Legge & Foley, 1980; Wichmann, 1999). This approach has not been without success: The contrast gain-control model, for instance, makes use of a narrowly tuned excitatory factor and a broadly tuned divisive inhibitory factor that give rise to a channel with a Mexican-hat-shaped weighting function (see Figure 4-2a) and explains much of the variance in sinusoidal contrast discrimination data (Foley, 1994; Wichmann, 1999). Further, the gain-control model correctly predicts some of the ways in which contrast-discrimination performance changes in the presence of an

additional sinusoidal masker having an orientation other than that of the signal (Foley, 1994).

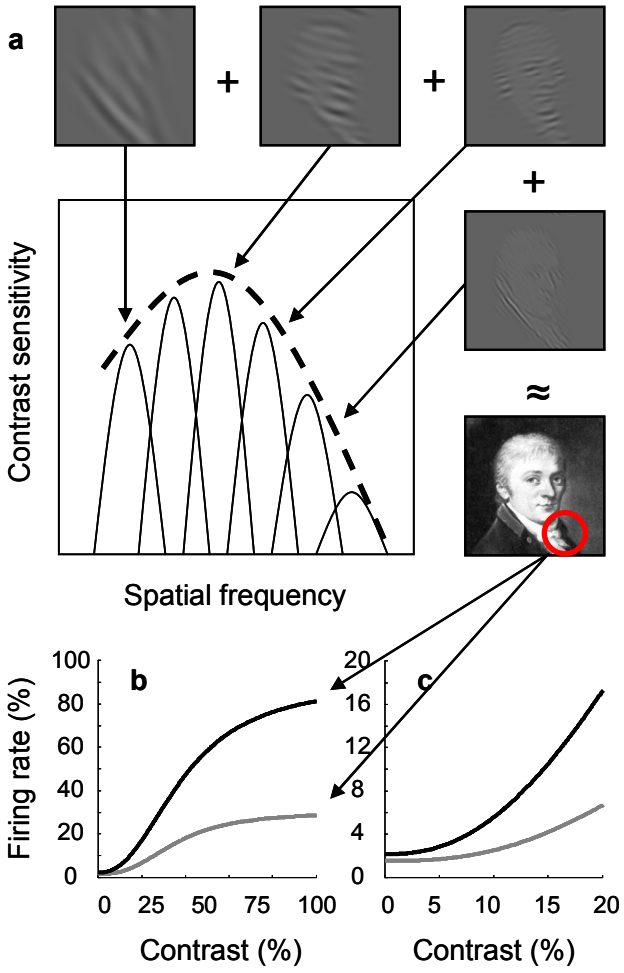

Figure 4-1. *The multi-channel model of spatial vision.* ***a.*** *Linear narrow-band filter responses to an example image. Particular filters' spatial-frequency tuning is shown in the main panel and their response illustrated in the surrounding panels. For simplicity in the illustration, the cells' orientation tuning has been ignored. Their weighted sum of responses approximately produces the characteristics of the full image. The dashed curve in the main panel of (a) depicts the whole contrast sensitivity function.* ***b.*** *The simulated contrast response functions of two hypothetical cortical neurons that behave as localised spatial-frequency and orientation-selective filters. Cell firing rate is plotted as a function of image contrast and the two filters are differentially sensitive to some narrow band stimulus. At all contrasts, the cell that is tuned to the spatial frequency and orientation spectrum of the image patch has a stronger response (upper dark line) than a less well tuned cell (lower grey line). Note that response saturation of both cells occurs at the same contrast.* ***c.*** *The same contrast response functions, expanded to show the response over the low-contrast range. At low contrasts, the contrast response function behaves as a positively accelerating non-linearity (a power law, see Finn et al., 2007; Miller & Troyer, 2002).*

However, the assumption that only the most sensitive channel is monitored has recently been challenged in a series of contrast-discrimination experiments (Henning & Wichmann, 2007): The pedestal effect is somewhat reduced in the presence of broadband noise or when either low-pass or high-pass noise distant in frequency from the spatial frequency of the signal is used. But when the high-pass and low-pass noises are combined to produce a "notched noise" from which a 1.5-octave wide notch centred on the signal frequency has been removed, the pedestal effect all but disappears. As illustrated in Figure 4-2, single-channel models like the contrast gain-control model fail to account for these results. Many different simulations of the gain-control model – more specifically, Foley's model 3 –, one of which is illustrated in Figure 4-2, all predict that the pedestal effect should be reduced in broadband noise, but persist in notched noise – the opposite of what in fact happens. This discrepancy may indicate that this implementation of divisive inhibition is not the appropriate computation to capture the effects of visual noise. However, in addition, there is a discrepancy between the deep dip attributed to single channels by gain-control models, and the mild dip observed in single cells of the striate cortex (Geisler & Albrecht, 1997). Thus, a single-channel explanation of the pedestal effect seems unlikely.

Bayesian models based on single-neuron characteristics that include the variability of the neurons' responses reconcile the mild dip in single cells and the strong pedestal effect observed psychophysically (Chirimuuta & Tolhurst, 2005). However, because these models do not consider the variability in the spatial-frequency tuning of cortical neurons, they would also fail to capture the contrast discrimination in noise results of Henning and Wichmann (2007). In sum, these results appear to require a revised model for human contrast discrimination.

Henning and Wichmann (2007) conclude that the pedestal effect may stem from the use of contrast information carried by channels tuned to spatial frequencies other than that of the signal frequency, thereby raising the important question of how the response of the whole population of channels is decoded in contrast discrimination and, indeed, in detection.

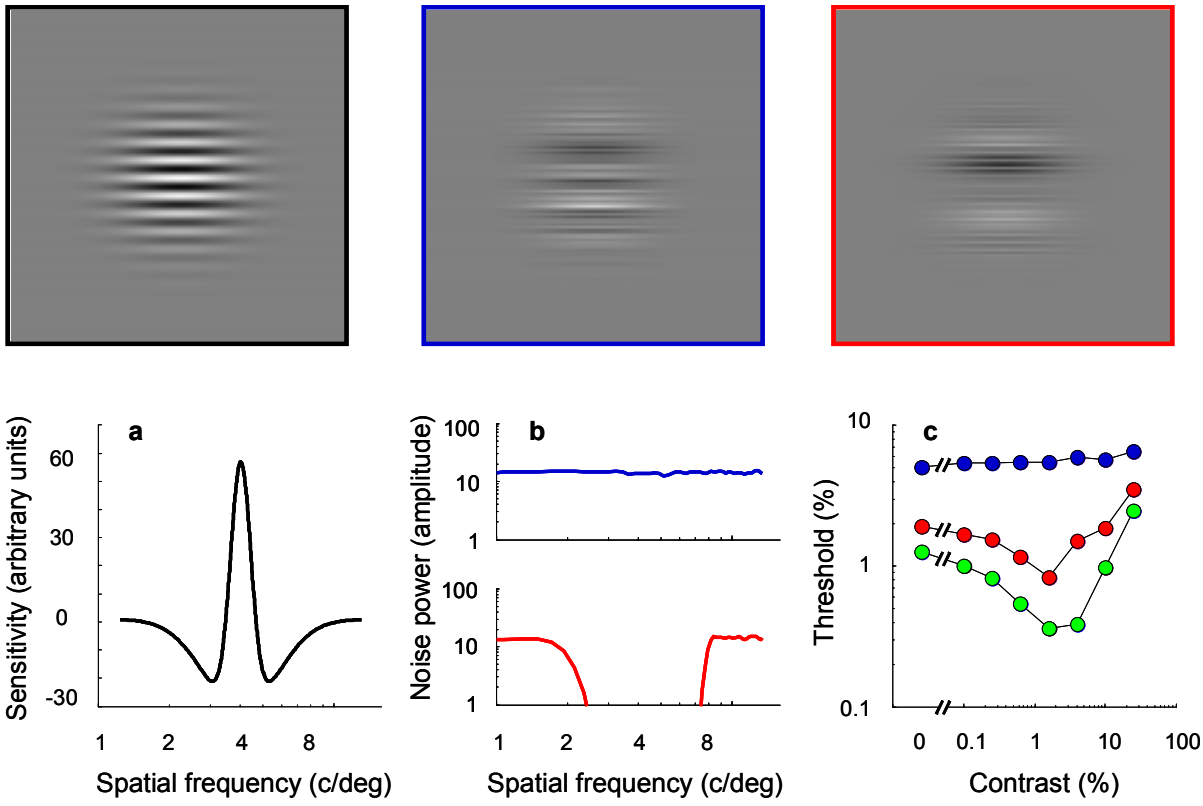

Figure 4-2. *Simulation of the contrast gain-control model (Foley's Model 3) in the contrast discrimination experiments of Henning and Wichmann (2007). Upper row: illustration of the signal and noise stimuli used in the simulation – from left to right: the 4 c/deg sinusoidal signal, broadband noise and notched noise. Lower row: a. The difference in contrast sensitivity of the excitatory and inhibitory channel components as a function of spatial frequency. The narrowly tuned excitatory component and broadly tuned inhibitory component give rise to a Mexican hat shaped channel. b. The noise-power densities as a function of spatial frequency – top: broadband noise (blue), bottom: notched noise (red). c. Performance predicted by the gain-control model in the contrast discrimination experiments of Henning and Wichmann (2007). Without noise added (green symbols), the model produces a dipper-shaped threshold-vs.-contrast function that closely mimics human data. Inconsistent with human performance, the dip is virtually absent in the presence of broadband noise (blue symbols). But the improvement in performance occurs in the presence of notched noise (red symbols), again inconsistent with human behaviour.*

In this paper, we evaluate the consequences for contrast-discrimination of simple weighted summation – where the contribution of each channel to detection or discrimination performance is proportional to its response. In particular, we consider possible neural correlates of "channels", where a channel is a set of neurons with similar selectivity for orientation and spatial frequency. Weighted summation of

signals from neural populations has been suggested to offer a general, biologically plausible mechanism capable of approximating observers' behaviour (Deneve, Latham & Pouget, 1999; Pouget, Zemel & Dayan, 2000). Further, under many realistic conditions, employing a simple response-based weighting heuristic even resembles optimal combination rules.

Here, we simulate a network made up of units whose contrast response functions resemble those found in cortex (see Methods). In our simulations we additionally use the population distributions of the units' characteristics even though information about population distributions is very limited. The units are restricted to physiologically measured characteristics to obtain a neurophysiologically plausible model. Because the neurophysiologically determined data for cortical cell responses to noise are very limited, our implementation of noise effects is necessarily based on theory-based predictions and simulations. Nevertheless, this implementation is consistent with available physiological findings and, as we shall see, the simulated networks predict plausible contrast-discrimination data. Further, contrary to the suggestion by Henning and Wichmann (2007), notched noise in our implementation does not prevent off-frequency looking, but modifies cortical cell activity in such a way that, upon pooling, the pedestal effect virtually disappears.

Neurophysiological recordings show that as the contrast of a sinusoidal grating driving cortical cells increases from zero, the cells' responses first increase in an accelerating fashion, remain approximately linear over a limited contrast range and then saturate (Albrecht, Geisler, Frazor & Crane, 2002; Geisler & Albrecht, 1997) (see Figure 4-1b,c). The saturation level is often attributed to some form of contrast-gain control. Both nonlinearities – response expansion at low contrasts and saturation at high contrasts – are usually thought to be fully expressed at the onset of the response and may play an important role in enhancing and maintaining stimulus selectivity (Albrecht et al., 2002; Geisler & Albrecht, 1997) although the dynamics of the physiological responses are not yet fully understood (Ringach, Hawken & Shapley, 1997).

Computationally, these properties are consistent with a model in which a strictly linear spatiotemporal stage is followed by squaring, then half-wave rectification, and broadband divisive inhibition – the normalization model (Carandini, Heeger & Movshon, 1997; Heeger, 1992a-b) –, though all the properties may be intrinsic to the feed-forward mechanisms of simple cells (Carandini, Heeger & Senn, 2002; Finn, Priebe & Ferster, 2007; Freeman, Durand, Kiper & Carandini, 2002; Miller & Troyer, 2002). The contrast-determined response saturation allows cortical neurons to signal information about the location, spatial frequency (or size), and orientation of local image features with considerable precision, at the expense of a detailed representation of contrast information (Albrecht et al., 2002; Geisler & Albrecht, 1997).

A final and particularly important property of cortical neurons is the characteristic that the variance of their responses to a sinusoidal grating is proportional to their mean response (Albrecht et al., 2002; Geisler & Albrecht, 1997; Vogels, Spileers & Orban, 1989). We call this crucial characteristic "the proportionality rule" and implement it in our network units as a multiplicative noise source. The simple proportionality relation implies that the ratio of a cell's mean response to the standard deviation of its responses – its signal-to-noise ratio – will increase as the response increases and this leads to the attractive idea that in combining the responses of many cells, the visual system would do well to first weight them by the strength of their response.

In the simulations to follow, we first discuss effects of unit selectivity and stimulus contrast on weight assignment in the model, making use of a pool of 100 units. We then consider, for the sake of clarity and simplicity, a situation where the pool of combined elements consists of only 12 uncorrelated units. One of these units is optimally tuned to the signal, while the tuning of the other units varies randomly from completely insensitive to highly sensitive. This somewhat artificial but illustrative situation enables an explicit comparison between the performance of an 'optimal' single unit and the performance of a network and thus allows an assessment of pooling effects. We show how this simple version of the model produces a pedestal effect based on informational pooling across spatial frequencies and is also able to

simulate successfully the findings of Henning and Wichmann (2007). We further consider grating *detection* (no pedestal) in noise and find that the model correctly predicts how detection performance changes in spectrally flat and filtered noise. Finally, having estimated the likely number of contributing units, we demonstrate that a more realistic pool consisting of 250 units with correlated responses and tuned to a broad range of spatial frequencies robustly produces similar results.

4.2 Methods

4.2.1 Model Equations

The contrast response functions of the units in the network simulations (Equation 4-1 below) are those of the *Invariant Response Descriptive Model* described in Albrecht et al. (2002), expanded to include an explicit selectivity parameter, *Sel*, which varies between 0 (i.e., the unit is not sensitive to the signal) and 1 (i.e., the peak sensitivity of the unit's spatial-weighting function corresponds to the spatial frequency of the signal). The selectivity parameter is needed for the units' response functions to have the behaviour demonstrated by cortical neurons (Albrecht et al., 2002; Geisler & Albrecht, 1997). The response functions, based on the Naka-Rushton equation, provide a good fit to the contrast response functions of striate cortex neurons to preferred ($Sel = 1$) and nonpreferred ($Sel < 1$) stimuli (Albrecht et al., 2002; Geisler & Albrecht, 1997). Equation 4-1 shows the mean response of a unit, $\overline{R_u}$, as a function of stimulus contrast c, expressed as a fraction of the unit's maximal firing rate:

$$\overline{R_u}(c) = r_0 + Sel\left(r_{\max}\frac{c^n}{c_{50}^n + c^n}\right), \tag{4-1}$$

where r_0 is a spontaneous discharge rate. In the simulations, r_0 is drawn from an exponential distribution with a mean value of 1.5% of the maximal firing rate (Olshausen & Field, 2005), [$r_0 \sim \mathrm{Exp}(1.5)$]; r_{max} is the maximum firing rate, drawn from a normal distribution with mean 81.8 and standard deviation 12.2, [$r_{max} \sim$

$N(81.8, 12.2)]$; n is the response exponent, $[n \sim N(2.4, 0.18)]$; c_{50} is the semi-saturation contrast, $[c_{50} \sim N(0.387, 0.0351)]$. (The expressions in square brackets following the definition of the terms in equations, give, where appropriate, the form and parameters of the distribution from which values for the terms were randomly selected. The parameter distributions are based on neurophysiologically determined estimates (Albrecht et al., 2002). However, the exact parameter settings are not critical to any of the claims made in the paper. Nevertheless, we shall see that these distributions produce a remarkably good approximation of psychophysical data.)

In the model, as in the measured behaviour of cortical cells, the variance of a unit's response is proportional to its mean value and is given by Equation 4-2:

$$Var(R_u(c)) = 1.5(\overline{R_u}(c)), \tag{4-2}$$

where $Var(R_u(c))$ is the variance of a unit's response as a function of stimulus contrast. The scaling value of 1.5 is based on estimates provided in several papers on cortical cell response reliability (Albrecht et al., 2002; Geisler & Albrecht, 1997; Vogels et al., 1989). The particular value of the proportionality constant is not critical, but the fact that variance is proportional to mean activity is crucial.

For the network simulations in this paper, the weight of each unit, $\omega_u(c)$, is fully determined by its mean response, normalised, for convenience, by the sum of all the contributing weights, and given by Equation 4-3:

$$\omega_u(c) = \frac{\overline{R_u}(c)}{\sum_{i=1}^{N} \overline{R_i}(c)}, \tag{4-3}$$

where $\omega_u(c)$ is the weight of unit u at stimulus contrast c in a network of N units. In order to make the simulations tractable, the trial-to-trial variation in the weights was ignored; we used the mean responses in calculating the weights. This is a simplification. In a real nervous system, the means would not, of course, be available

from a single unit and weights would necessarily be based on responses alone. In Appendix 4-1, we demonstrate that this simplification is immaterial with respect to the conclusions we draw.

The mean pooled response at any given contrast, $\overline{R_{pooled}}(c)$, is:

$$\overline{R_{pooled}}(c) = \sum_{u=1}^{N} \omega_u(c)\overline{R_u}(c)\,. \tag{4-4}$$

The more responsive and hence, because of the proportionality rule, the more reliable units thus attract more weight. It will be noted in passing that this is in effect a crude, heuristic approximation to an optimal decoder for discriminating multivariate uncorrelated signals (Strang, 1986) – an idealised and simplified description of our network. An optimal decoder would weight the units' contrast estimates by the inverse of their covariance matrix, which, if the units were uncorrelated, would lead to weights that were proportional to the reciprocal of the variances. We do not wish to suggest that the decoding is necessarily as simple as our heuristic rule. But its simplicity is appealing and for most situations, it resembles an optimal combination rule without the computational burden of determining the covariance matrix for the units in the network or, indeed, knowing anything about the precision of any unit – because the variance is proportional to the responsiveness, only the strength of the response to any given stimulus matters.

Finally, to compute the variance of the pooled responses, we made use of the pooling formula (Shadlen & Newsome, 1998):

$$Var(R_{pooled}(c)) = \frac{1}{N^2}\left(\sum_{u=1}^{N}\omega_u^2(c)Var(R_u(c)) + \sum_{u=1}^{N}\sum_{v\neq u}^{N} r_{uv}\sqrt{\omega_u^2(c)Var(R_u(c))\omega_v^2(c)Var(R_v(c))}\right), \tag{4-5}$$

where r_{uv} is the correlation coefficient between the u^{th} and the v^{th} unit. Correlation among units is, of course, notoriously difficult to determine. Nevertheless, it is an

important – indeed a crucial – factor in some models of MT pooling (Shadlen, Britten, Newsome and Movshon, 1996).

Contrast discrimination performance of the network, expressed as d', is fully determined by the mean and variances of the pooled responses to the pedestal and to the pedestal-plus-signal (Green & Swets, 1966) and given by Equation 4-6):

$$d'_{pooled}(c_{ped}, c_{ped+sig}) = \frac{\overline{R_{pooled}}(c_{ped+sig}) - \overline{R_{pooled}}(c_{ped})}{\sqrt{Var(R_{pooled}(c_{ped+sig})) + Var(R_{pooled}(c_{ped}))}}, \quad (4\text{-}6)$$

4.2.2 Cortical neurons and visual noise

Early neurophysiological work demonstrated some heterogenity among cortical neurons in cat striate cortex in response to noise. Simple cells have been reported to be unresponsive to a broadband noise stimulus while their response to an otherwise optimal stimulus is inhibited when this stimulus is embedded in noise (Hammond & MacKay, 1977; Maffei, Morrone, Pirchio & Sandini, 1979). Other neurons (mainly complex cells) have been reported to respond to (some kinds of) broadband noise, though their response to an otherwise optimal stimulus is also inhibited when this stimulus is embedded in noise (Hammond & MacKay, 1977; Maffei et al., 1979). For both kinds of neurons, the main effect of spectrally flat noise is thus that single units' contrast-response functions shift toward higher contrasts and lower response rates.

More recent work by Carandini et al. (1997) demonstrated that the effects of binary noise on the contrast-response functions of simple cells in macaque primary visual cortex are well captured by the normalization model mentioned briefly in the introduction. In their paper, Carandini et al. (1997) make the simplifying assumption that the noise would be unable to drive the linear receptive field of the cells, so that the sole effect of the noise would be to provide divisive normalization. To fit their data, Carandini et al. (1997) introduced an additional parameter α, controlling the effectiveness of noise in driving the normalization pool and reported that the values of α resulting from the fits to 22 simple cells were equally spread (on logarithmic

coordinates) between 0.1 and 10, indicating that the noise provided very strong inhibition for some cells, but only weak inhibition for others.

In this paper, we largely follow the implementation of Carandini et al. (1997) to capture the effects of noise on units in our model. However, we do not assume that broadband noise is unable to drive the linear receptive field for several reasons. First, Carandini et al. reported a 3-fold elevation of the maintained discharge in noise. This is unlikely to be explained by suppression and implies activation; Second, noise activation is crucial to explain some psychophysical observations –contrast detection performance improves slightly when weak noise is added to a low-contrast signal (Goris, Zaenen & Wagemans, 2008); Third, in the logic of the normalization model, the crucial difference between notched noise and flat noise for cells tuned to frequencies in the notch is that notched noise provides only non-specific suppression, while flat noise provides both excitatory and inhibitory activation.

The equations used to describe unit responses in broadband and notched noise are derived from Carandini's implementation combined with simulations and explained in detail in Appendix 4-2 (Equations 4-A.3 and 4-A.4). In summary, in the model developed here, addition of broadband noise increases the maintained discharge, modifies the response exponent and leads to an effective shift of the contrast response function toward higher contrasts and lower response rates. The strength of this shift varies across units, as has been reported for simple cells.

On the other hand, while the addition of notched noise produces similar effects, its effect is modified by the presence of the notch and by units' selectivity to frequencies both within and outside the notch. For units tuned to the frequencies outside the notch, notched noise is effectively equivalent to white noise, but for units tuned to frequencies in the notch, notched noise produces strong non-specific suppression. This is illustrated in Figure 4-3. Figure 4-3a shows how the mean response of one particular unit – tuned to the signal frequency in the centre of the notch – varies with signal contrast when different levels of notched noise are added to the signal. As the noise power increases (the more reddish curves), the response of the broadly

tuned normalization pool increases and becomes more variable, which suppresses and effectively linearizes the contrast response function. This is shown in Figure 4-3b. The response functions shown in panel a are replotted in panel b, normalized by the response at maximal signal contrast. Note that the response acceleration decreases with noise power. However, in the presence of notched noise, it is not this unit, as we shall see, that attracts the greatest weighting in the pool.

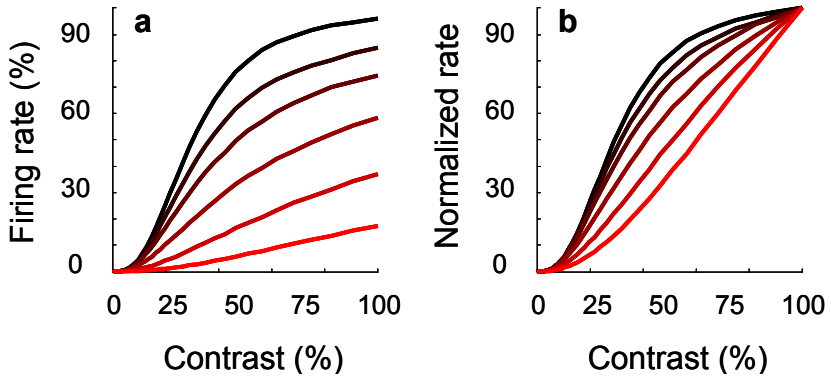

Figure 4-3. *The effects of notched noise.* ***a*** *The contrast response function for one particular unit – tuned to the centre frequency in the notch, with* $r_0 = 0$, $n = 2.6$ *and* $c_{50} = 0.3$ *– as a function of signal contrast with different levels of notched noise added to the signal. As the noise power increases (the more reddish curves), the response of the broadly tuned normalization pool increases and becomes more variable, which suppresses and effectively linearizes the contrast response function.* ***b.*** *The same response functions as in (a), normalized by their respective response at maximal signal contrast.*

The characteristics of the noise effects are based on a relatively small number of physiological observations combined with theory-based simulations, thus it is at best an approximation. Nevertheless, the simulated networks robustly predict plausible contrast-discrimination data, as we shall see.

4.3 Results

In the simulations described below, we first discuss effects of unit selectivity and stimulus contrast on weight assignment in the model, making use of a pool of 100 units. We then consider a situation where the pool consists of 12 uncorrelated units.

We show how this simple version of the model produces a pedestal effect based on weighted pooling across spatial frequencies and successfully simulates the findings of Henning and Wichmann (2007). Finally, we demonstrate that a more realistic version of the model produces similar results.

4.3.1 Weight assignment: effects of unit selectivity and stimulus contrast

In our network, each element's contribution depends simply on the magnitude of its response. Since the reliability of the unit is proportional to the mean response, such a weighting provides a crude approximation to optimal weighting based on the inverse of the covariance matrix for uncorrelated units (Strang, 1986). Figure 4-4a illustrates response rate as a function of contrast for a family of cortical neurons responding, in the absence of added external noise, to a narrow-band stimulus at their preferred orientation (Albrecht et al., 2002; Geisler & Albrecht, 1997).

In Figure 4-4a, the units with high saturation levels are those best tuned to the spatial frequency of the stimulus while progressively lower saturation levels represent units that respond less and less to the stimulus because their optimal stimuli have spatial frequencies that are further and further away from the spatial frequency of the stimulus. The close relation between saturation level and selectivity for spatial frequency allows us to use the relative saturation level of each unit as an indication of its selectivity for the sinusoidal signal and pedestal (see Equation 4-1). The saturation level of the 100 units in this simulation ranged from 100% to 1% of the maximum saturation level – a range that would represent a range of peak-sensitivities of roughly 5 octaves, uniformly sampled on a logarithmic axis, around the frequency of the signal (we assume geometrically symmetrical tuning functions with a bandwidth of approximately 1.5 octaves). These selectivities were mainly chosen for convenience. (Note that in behavioural studies, the contrast of narrowband stimuli rarely exceeds 50% and even this may be an order of magnitude greater than the contrasts that typically occur naturally (Frazor & Geisler, 2006). Thus the behaviourally relevant region of Figure 4-4a lies below about 30% contrast.)

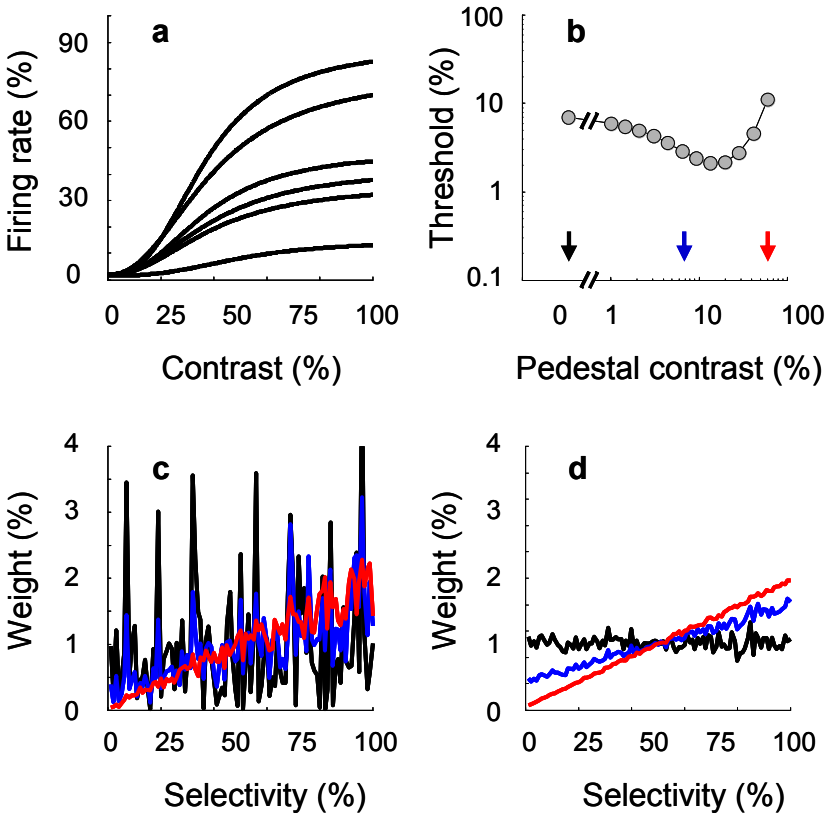

Figure 4-4. *Illustration of the effects of stimulus contrast on the distribution of weights. In this simulation, the network consisted of 100 units, with stimulus selectivity indicated by the relative saturation levels varying from 1% to 100% in steps of 1%.* ***a.*** *The contrast response functions to some narrow-band stimulus of six of the network units with similar orientation tuning but varying spatial-frequency tuning.* ***b.*** *The dipper-shaped threshold-vs.-contrast function of the network at 75% correct.* ***c.*** *Weight as a function of stimulus selectivity at the three different pedestal contrasts indicated by the arrows in b.* ***d.*** *Same as c, but averaged over 100 simulations.*

The behaviour shown in Figure 4-4b captures the central features of contrast discrimination that appear in experiments without added external noise. Figure 4-4b illustrates the network's dipper-shaped threshold-vs.-contrast function at 75% correct in a 2AFC task; the "dipper" shape constitutes the pedestal effect. Almost no unit in the simulation shows the pedestal effect by itself, rather the effect emerges from the pooled responses as will be discussed below.

Figures 4-4c and 4-4d both show the fraction of the total weight assigned to different units as a function of their selectivity. Results for three different pedestal contrasts

are shown: zero contrast in black, a just-visible pedestal contrast in blue, and a highly-visible pedestal contrast in red. These contrasts are indicated by the coloured arrows in Figure 4-4b. Figure 4-4c shows the distribution of weights as a function of unit selectivity for one simulation; Figure 4-4d shows the weights averaged over 100 simulations. For each of these simulations, the set of cell selectivities was held constant, but all other parameters of our network were randomly selected from the distributions described in the Methods section.

When a blank stimulus – a uniform field of the same mean luminance as the sinusoidal grating – is presented (black curves) the units by definition respond at their (usually nonzero) spontaneous discharge rate. Consequently, the jagged black line reflects the probability distribution of spontaneous activity across cells. The distribution was chosen to be exponential (Olshausen & Field, 2005) with an average of 1.5% of maximal firing rate (see Methods). This cell characteristic – base rate activity – is independent of stimulus selectivity and the distribution of weights is uniform with respect to selectivity (as can be seen in Figure 4-4d where the black line is the average weight in 100 simulations and is approximately horizontal).

As stimulus contrast increases from zero, the relative magnitude of a unit's response begins to reflect not only its spontaneous rate but also a stimulus driven part – the latter determined by the unit's selectivity (see Equation 4-1). Hence, at nonzero stimulus contrasts, higher weights are attracted by the better tuned units (i.e., more responsive units) in a fashion that is roughly linearly related to selectivity (the blue and red lines in Figure 4-4c and d). The higher weighting occurs, of course, because the weighting of a unit increases as its response increases. However, both the slope and variability of this linear relation depend on stimulus contrast.

At low, barely visible pedestal contrasts (blue curves) spontaneous activity contributes a high proportion of the total cell response. As a consequence of the exponentially distributed spontaneous discharge levels, a few units with high base-rate activity will attract much of the weight (the blue curve in Figure 4-4c). Note that such units would be called 'irrelevant' in the context of uncertainty models

(Pelli, 1985). As pedestal contrast increases, more and more cells are driven by the stimulus and spontaneous activity gradually looses its influence on network activity. Consequently, the variability across the distribution of weights decreases with increasing contrast (the red curve in Figure 4-4c) and the slope of the distribution of weights increases (the red curve in Figure 4-4d).

Thus, with contrast detection and discrimination based on the combined response of many elements instead of a single element, the class of units contributing most to the decision statistic and distinguished by their stimulus selectivity or spatial-frequency tuning, varies a lot. With weighting determined by responsiveness and especially at low stimulus contrasts, units tuned to spatial frequencies and orientations remote from the signal are often weighted heavily.

We now investigate in more detail the effects of weighted pooling on contrast detection and discrimination. In the following simulations, we consider the simplified situation where the pool consists of 12 uncorrelated units. One of these units is optimally tuned to the signal (i.e., selectivity = 1), and the selectivity of the 11 other units, as indexed by their saturation level, is randomly drawn from a Gaussian distribution centred at 0.50, with a standard deviation of 0.17 and clipped at 0 and 1. For the 4-c/deg signal used in our simulations, this selectivity distribution is not inconsistent with spatial frequency tuning properties of cortical neurons (Geisler & Albrecht, 1997); i.e., many neurons are somewhat sensitive to the signal, while few are optimally tuned to the signal, or completely insensitive to the signal.

4.3.2 Contrast discrimination: effects of response pooling

Figure 4-5a illustrates response rate as a function of contrast for all twelve units in one particular pool driven, in the absence of added external noise, by a narrow-band stimulus at their preferred orientation. The contrast response function of the unit that is optimally tuned to the signal is shown in purple. In comparison to monitoring only the most sensitive unit, pooling the responses from neurons of different sensitivity weighted by their responsiveness, i.e., weighted pooling across units,

improves detectability at all stimulus contrasts except in some cases where the contrast is so low that few units are stimulus driven. Moreover, because of the proportionality rule, the improvement in sensitivity from weighted pooling relative to the most selective single unit increases as contrast increases.

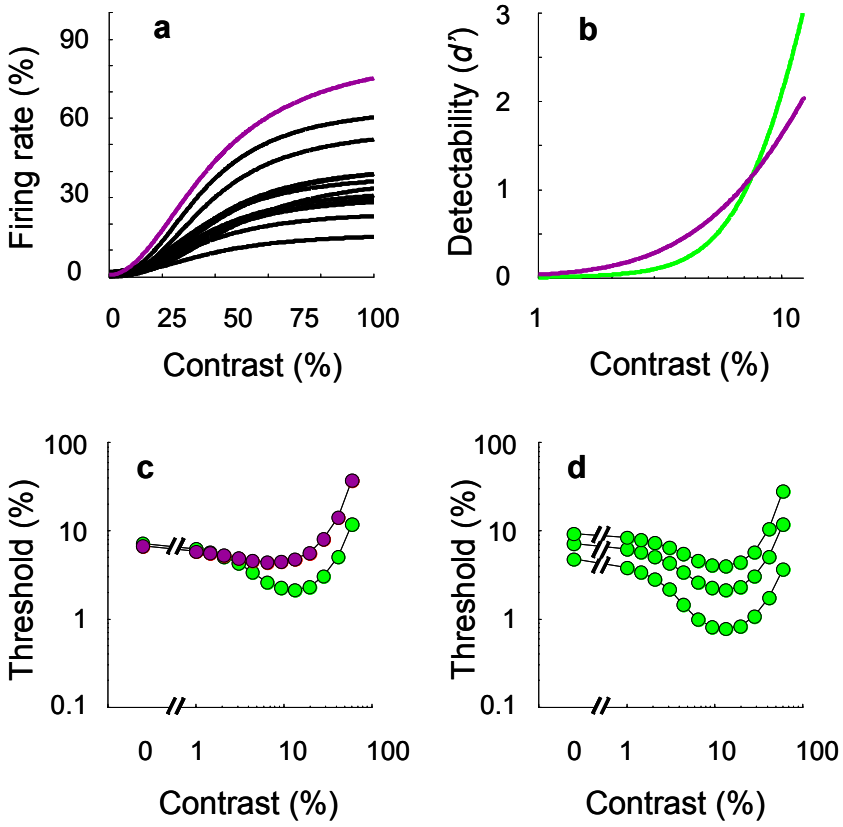

Figure 4-5. *Illustration of the effects of pooling on contrast discrimination.* ***a.*** *The response functions of the twelve units in one particular pool to signals of increasing contrast. The contrast response function of the unit that is optimally tuned to the signal is shown in purple.* ***b.*** *Detectability, d', expressed as the ratio of the mean response to the standard deviation, as a function of contrast for the most sensitive unit (purple line) and the pooled decision statistic (green line) on semi-logarithmic coordinates.* ***c.*** *Contrast discrimination thresholds at 75% correct as a function of pedestal contrast on double logarithmic coordinates for the most sensitive unit (purple) and the pooled decision statistic (green).* ***d.*** *Contrast discrimination thresholds as a function of pedestal contrast on double logarithmic coordinates for the pooled decision statistic at three performance levels in 2AFC (lowest row: 60%, middle row: 75% and upper row: 90%).*

To see this consider Figure 4-5b, which illustrates how the ratio of the mean pooled response to the standard deviation of the pooled response depends on the contrast of a narrowband stimulus. This ratio is closely related to the detectability of the stimulus and is labelled d' (Green & Swets, 1966). The ratio of the mean to the standard deviation for the pooled group is indicated by the green line, the same ratio for the most selective unit, by the purple line. The higher this ratio, the better the system is able to discriminate a low-contrast signal grating from a uniform field. At the very lowest stimulus contrasts, where the non-optimally tuned units mainly contribute noise, the purple line lies above the green line, indicating that performance of the most selective unit is slightly better than that of the pooled response. From a certain stimulus contrast on, however, the green line lies above the purple line, indicating that the pooled response outperforms the most selective unit. Moreover, as contrast increases, the difference between the two functions grows as a consequence of the proportionality rule. The pooled detectability function is thus more sharply accelerated and it has been suggested that this particular nonlinearity underlies the pedestal effect (Nachmias, 1981; Nachmias & Sansbury, 1974; Smithson, Henning, MacLeod and Stockman, 2009).

Figure 4-5c shows the 75% correct "thresholds" for detecting a narrowband stimulus as a function of the contrast of the pedestal, again in the absence of noise; the threshold for the pooled group is indicated by the green symbols, for the most selective unit, by the purple. In both cases the thresholds are determined by calculating the difference between the mean response to the pedestal alone and the mean response to the signal-plus pedestal and dividing this difference by the square root of the sum of their variances (see Equation 4-6). It is clear in Figure 4-5c that the pooling of information in the network leads not only to better performance from a certain pedestal contrast on, but also to a bigger pedestal effect. The pedestal effect of the pooled network is based on broadband weighted pooling, i.e., on the use of information carried by units tuned to frequencies other than the signal frequency – i.e. a form of off-frequency looking. The finding that pooling responses weighted by their responsiveness increases the size of the pedestal effect, in addition, removes the puzzling discrepancy between the mild dip observed in single cells in striate cortex

(Geisler & Albrecht, 1997), and the deep dip observed psychophysically (Nachmias & Sansbury, 1974; Bird et al., 2002; see also Chirimuuta & Tolhurst, 2005).

In human vision, the strength of the pedestal effect is known to depend on the performance level taken to define the "threshold" (Henning & Wichmann, 2007; Nachmias & Sansbury, 1974; Bird et al., 2002; Goris, Wagemans & Wichmann, 2008; Wichmann, 1999). At low performance levels, the maximal pedestal-induced threshold reduction is considerably larger than at high performance levels. This particular property of contrast discrimination has largely been ignored because only one performance contour is usually determined (Foley, 1994; Legge & Foley, 1980) but has already proven useful in model selection (Goris, Wagemans & Wichmann, 2008; Wichmann, 1999). In Figure 4-5d, discrimination thresholds for the pooled network of weighted unit activation are plotted at 60%, 75% and 90% correct response levels. The network clearly gives rise to the observed performance-level-dependent pedestal effect: Consistent with human data, the dip of the threshold-vs.-pedestal contrast function is deeper at lower performance levels.

Another salient feature of contrast-discrimination studies is that, from a certain pedestal contrast level on, discrimination thresholds as a function of contrast plotted on double logarithmic co-ordinates rise in a nearly linear fashion (e.g., Bird et al., 2002). This, of course, is Weber's law. It can be seen in Figure 4-5c that the rise in discrimination thresholds of the pooled population response is also approximately linear on double logarithmic coordinates. This is a consequence of the roughly constant ratio of the variance and mean of the pooled response in the high-contrast region. Thus, because the variance of the pooled response is effectively proportional to the mean pooled response at suprathreshold contrasts, the ratio of the just-noticeable contrast increment and pedestal contrast is approximately constant, consistent with Weber's Law and the linear slope on double logarithmic co-ordinates.

The slope of this part of the dipper curve on the co-ordinates of Figure 4-5c is steeper than psychophysical estimates that are typically around or slightly below one (Bird et al., 2002; Wichmann, 1999). However, if one or more of the network units

saturated more slowly at higher contrasts, responding more linearly over the entire contrast range, the slope would be lower. Indeed, there is a great deal of heterogeinity among cortical cells and about 5% demonstrate a nearly linear relation between contrast and response magnitude (Albrecht et al., 2002).

The effect of reaching saturation more slowly, i.e., extended linearity may be twofold: First, because their responses continue to increase after other units saturate, the more linear units will attract an increasing proportion of the weighting, provided, of course, that such units ultimately exhibit high saturation levels. Thus, although they may constitute a small proportion of the population, their contribution to the pool will be disproportionately large at high contrasts. Second, because of the large contribution of the linear units at high contrasts, the pool will be effectively linearized. The importance of linearization at higher levels of the visual system is a major issue (Eliasmith & Anderson, 2003). (See Chirimuuta and Tolhurst (2005) for an alternative solution to fit the rising part of the dipper function).

In summary, pooling the responses of a limited number of units that resemble the contrast response functions and statistical properties found in primary visual cortex by simple weighted summation produces contrast discrimination predictions that resemble several features of human vision. Further, the pedestal effect in this model is based on the use of information carried by units tuned to frequencies other than the signal frequency. This is, of course, a form of off-frequency looking.

We now specifically investigate whether this simple implementation of off-frequency looking is able to capture the contrast discrimination in noise results reported by Henning and Wichmann (2007).

4.3.3 Contrast discrimination in noise: effects of response pooling

Thus far, we have considered contrast processing in the absence of added external noise. Addition of spectrally flat noise to a narrow-band stimulus increases cell activity at low response rates, but inhibits cell activity at high response rates

(Carandini et al., 1997; see Appendix 4-2). Presumably, the activation is caused by rectification and the inhibition by broadband normalization. These effects are thus in line with the normalization model, and – as explained in detail in Appendix 4-2 – can be captured by extending the descriptive equation of the contrast response function (Equation 4-A.3). Finally, because noise elicits variable excitatory and inhibitory activation, the acceleration of the contrast response function is softened by noise (Miller & Troyer, 2002).

To the best of our knowledge, no physiological data on effects of notched noise are available. We thus base our predictions on the logic of the normalization model. For a unit tuned to the frequency in the centre of the 1.5-octaves wide notch, notched noise can only provide non-specific suppression. This is because the noise will have a negligible effect on spatial-frequency tuned excitatory frequencies of the unit. Given that broadband noise increases cell activity at low response rates, the inhibitory effect of the notched noise is likely to be stronger than the net inhibition provided by the broadband noise. In our simulations, notched noise is assumed to cause twice as much inhibition as white noise for a unit tuned to the signal frequency (see Equation 4-A.4 in Appendix 4-2). For a unit tuned to a frequency outside the notch however, the effects of the noise are assumed to be approximately similar to effects of broadband noise. As was the case for broadband noise, the acceleration of the contrast response function is softened by notched noise. Importantly, simulations with the normalization model showed that stronger noise suppression produces a more linear contrast response function (Figure 4-3). Thus the linearizing effects of notched noise are most pronounced for units tuned to the signal frequency (i.e., Sel = 1) and decrease with selectivity.

In summary, the effects of broadband noise and notched noise on the contrast response function were derived from model simulations and described by Equation 4-A.3 and 4-A.4. Our model has only one free parameter, i.e., the effective noise power. At a noise power which produces on average a four-fold elevation of the maintained discharge, we find that the model robustly produces contrast-discrimination data taken in noise, that are consistent with the observations of Henning and Wichmann

(2007): the pedestal effect is slightly diminished in broadband noise but disappears almost completely in notched noise. (This noise power is not unreasonably large: Carandini et al. (1997) – making use of binary noise at several noise powers – report an average three-fold elevation in their experiments; Henning and Wichmann (2007) used the maximum noise-power density (1D noise) their display system allowed them to generate.)

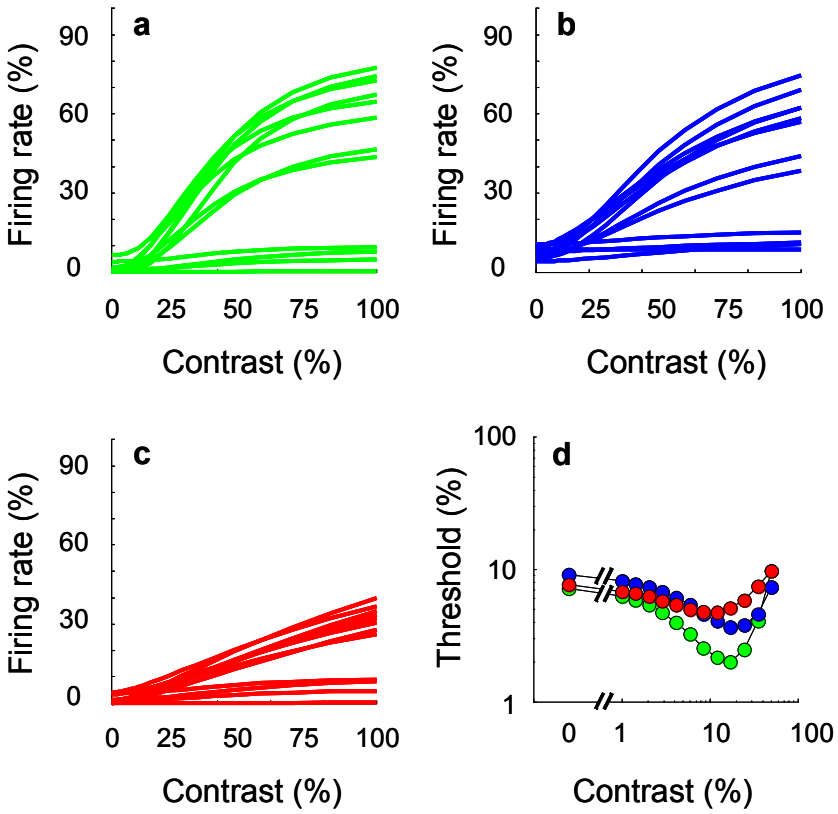

Figure 4-6. *Contrast discrimination in noise.* ***a.*** *The contrast response functions of the twelve units in one particular pool to the signal with no added external noise.* ***b.*** *The contrast response functions for the signal embedded in broadband external noise.* ***c.*** *The contrast response functions for the signal in the presence of notched noise.* ***d.*** *Predicted contrast discrimination thresholds for this pool as a function of pedestal contrast with no external noise (green symbols), in broadband noise (blue symbols) and in notched noise (red symbols).*

Figure 4-6 illustrates the twelve noiseless contrast response functions for one particular pool with no external noise (Figure 4-6a), as well as the changes these functions undergo in spectrally flat (Figure 4-6b) and notched noise (Figure 4-6c).

The contrast-discrimination functions for each of these conditions are shown in Figure 4-6d. The contrast-discrimination-threshold functions shown in Figure 4-6d closely mimic the results of Henning and Wichmann (2007): the pedestal effect in the absence of noise (green) is shifted on these co-ordinates to higher contrasts in broadband noise (blue) and slightly diminished, but disappears almost completely in notched noise (red). Further, at higher pedestal contrast levels, discrimination thresholds of all noise conditions almost coincide, consistent with human data (Henning & Wichmann, 2007). We thus conclude that our implementation of noise effects and broadband weighted response pooling is able to produce plausible results for contrast discrimination in noise.

4.3.4 Contrast detection: effects of response pooling

Pooling the responses of many units with wide heterogeneity in spatial-frequency tuning produces a detectability function that is largely shifted to lower contrasts and more sharply accelerated relative to the detectability function for a single-unit (see Figure 4-5b). The model presented in this paper thus shows that response pooling is a sensible strategy for contrast detection as well. Consequently, experimental manipulations that modify the contrast response functions of cortical cells will affect detection performance. Although not discussed by Henning and Wichmann (2007), the detection data they also gathered in their series of contrast-discrimination experiments provide a test of this hypothesis.

In their experiments, detection of a 4.0-c/deg sinusoidal grating was measured without noise and with notched noise added – the notch being 1.5 octaves wide and geometrically centred on the signal frequency. Figure 4-7a shows how signal detectability increases as a function of contrast without noise (green), in broadband noise (blue) and in notched noise (red) according to the model. It will be noted that these functions are closely related to the strength of the pedestal effect, i.e., more rapidly accelerating detectability functions correspond to a stronger pedestal effect. Figure 4-7b plots the predicted psychometric functions, i.e., percentage correct in a 2AFC-task as a function of signal contrast on semi-logarithmic coordinates, derived

from the detectability functions shown in Figure 4-7a. First consider detection without noise and in notched noise.

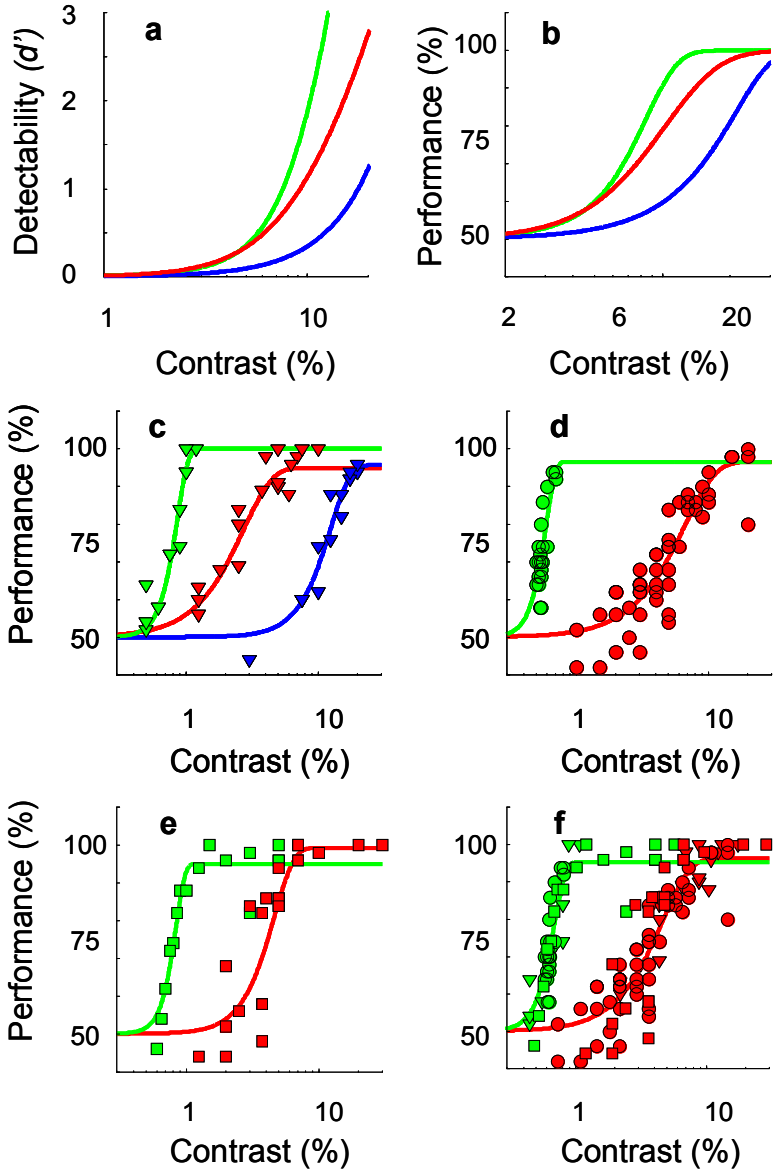

Figure 4-7. *Illustration of the effect of channel pooling on contrast detection.* ***a.*** *Signal detectability as a function of contrast for the pooled decision statistic on semi-logarithmic coordinates. Performance without noise is shown in green, in notched noise in red and in white noise in blue.* ***b.*** *The psychometric functions relating percentage correct in a 2AFC detection task to signal contrast (corresponding to the detectability functions shown in panel a).* ***c.*** *Detection performance in a 2AFC-task as a function of signal contrast for observer GBH. Green symbols refer to the no-noise condition, red symbols to the notched-noise condition, blue symbols to the broadband-noise condition.* ***d.*** *Detection performance for observer NAL.* ***e.*** *Detection performance for observer TCC.* ***f.*** *Detection performance as a function of (rescaled) signal contrast for all observers in the no-noise (green) and notched-noise (red) conditions.*

According to the model, detection performance without noise should be better at most stimulus contrasts (see Figure 4-7b). Furthermore, the psychometric function relating performance (% correct) to signal contrast should be steeper without notched noise. Figure 4-7c-e shows the percentage of correct responses as a function of signal contrast on semi-logarithmic coordinates obtained separately from three observers with no noise (green) and with (red) notched noise (Henning & Wichmann, 2007). For observer GBH, detection data with broadband noise (blue) are plotted as well. Fitted psychometric functions are also shown.

As can be seen, the addition of notched noise not only hurts detection performance, but also produces a more shallow psychometric function. Figure 4-7f shows the psychometric functions fitted to the combined data of all three observers, where the different symbols refer to different observers. (For each of the relevant conditions, the data for each observer have been scaled to equate the observers' 75% correct level within a given condition, thus increasing the statistical power of the test designed to compare the slopes of the psychometric functions.) It is clear that the data are consistent with the hypotheses formulated above: The signal contrast necessary to achieve 75% correct is lower in the no-noise condition (for each observer, $p < 10^{-6}$), and analysis of the β-parameter – the parameter controlling the steepness of the cumulative Weibull function fitted to the data – reveals that the psychometric function is also steeper than for the notched noise condition (for each observer, $p < 0.05$, for the combined data, $p < 10^{-6}$) (Wichmann & Hill, 2001). These data are thus consistent with the notion that even detection of a sinusoidal grating may be based on pooled responses rather than on the most responsive channel.

For the model, the detection psychometric function in broadband noise is shallower than without noise, but steeper than in notched noise if the noise-power density in the pass bands is kept constant. The 75% correct threshold, on the other hand, is expected to be higher in white noise than in both other conditions. Henning and Wichmann (2007) also measured detection in white noise, but noise power in the pass bands was only kept constant for observer GBH. Nevertheless, it is interesting to note that his data suggest that the detection psychometric function in white noise is

indeed shallower than without noise ($p < 0.01$), but steeper than in notched noise ($p < 0.05$). Furthermore, as can be seen in Figure 4-7c, the signal contrast necessary to achieve 75% correct is higher in the white-noise condition than in the no-noise condition ($p < 10^{-6}$) and the notched-noise condition ($p < 10^{-6}$). These data are thus consistent with the notion that noise may modify the neurons' contrast response functions, without, however, altering the pooling rules.

In summary, the population code model that successfully simulates contrast discrimination in noise was shown to predict effects of broadband weighted pooling in detection as well. These predictions were tested and confirmed, thus suggesting that even the detection of a low-contrast sinusoidal grating may be based on the responses of many elements with a wide heterogeneity in spatial frequency tuning.

4.3.5 Larger pools with random spatial frequency tuning and correlated noise

A final consideration is the issue whether this model is able to produce realistic results in a less artificial situation. In the following simulation, we therefore consider a pool consisting of 250 correlated units tuned to a broad range of spatial frequencies.

In the type of network discussed in this paper, where the main noise source is signal-dependent, allowing more units to contribute to the decision statistic leads to better detection performance and a stronger pedestal effect. At first sight, larger pools may thus be expected to produce stronger pedestal effects. However, in this regard it is important to note that single cell recordings have demonstrated that the responses of different cortical neurons in discrimination tasks are typically weakly correlated (Golledge et al., 2003; Panzeri, Golledge, Zheng, Tovée & Young, 2001; Montani, Kohn, Smith & Schultz, 2007; Shadlen et al., 1996; Zohary, Shadlen & Newsome, 1994). It is thus likely that channel responses are correlated to some extent. Indeed, at the psychophysical level, experimental evidence suggests that spatial-frequency channels may share some of their internal noise in contrast discrimination (Henning et al., 2002).

Correlated noise has two fundamentally different effects (Averbeck, Latham & Pouget, 2006): First, if unit activity is pooled according to a weighted average rule and, as proposed here, all weights are positive, correlated noise will substantially decrease the encoding capacity of the pool. Further, when the responses of units are correlated, pooling improvement with increasing numbers of units contributing to the pool approaches a limit once the number of units exceeds some critical number (Zohary et al., 1994). Second, correlated units might influence the computational decoding strategies appropriate for networks of neurons (Abbott & Dayan, 1999; Chen, Geisler & Seidemann, 2006). If much additional information can be gained by taking into account the fact that neural noise is correlated, the decoding strategies used in the brain may be affected, i.e., it may be appropriate to introduce a whitening stage.

For primary visual cortex (Golledge et al., 2003; Panzeri et al., 2001) correlations are estimated to be on the order of 0.1 to 0.15, but, Golledge et al. (2003), making use of information theory techniques to quantify the role of such small correlations, argue they would contribute less than 10% extra Shannon information in encoding visual information and it has been argued that such small correlations are unlikely to be taken into account in the decoding computations (Averbeck et al., 2006). Consequently, we do not consider whitening for our network.

Figure 4-8a illustrates the effects of correlated noise in our network. Detectability for a low-contrast grating (no pedestal) is plotted in Figure 4-8a as a function of the number of pooled units. Detectability is measured as the ratio of the mean response to its standard deviation, expressed as d'. Exactly one unit in the pool was optimally tuned to the signal. The selectivities of all other units were again randomly chosen from a Gaussian distribution centred at 0.50 with a standard deviation of 0.17 and clipped at 0 and 1. The average inter-unit correlation is coded by the colour indicated by the colour bar on the right of the figure. If the noise is uncorrelated [the highest (red) curve], addition of more units progressively improves detectability and, in the limit, would yield an errorless observer. On the other hand, weakly correlated noise (the other curves) show that the addition of more units has very little effect on

detectability once a certain critical number of units is reached. As the average inter-unit correlation increases (downward or more blue in Figure 4-8a), that critical number of units drops. For the degree of inter-neuron correlation between 0.1 and 0.2 (typical estimates from single cell recordings) the improvement with increasing numbers approaches its asymptotic level from 50 to 100 units (Zohary et al., 1994). Pool size is thus expected to be rather limited.

In the following simulation, we consider pools of 250 units – thus well above the critical number – with an average inter-neuron correlation of 0.15 (the standard deviation of this correlation equals 0.05). We further postulate a spatial-frequency tuning function for each unit. To approximate tuning functions of cortical neurons (see e.g. Geisler & Albrecht, 1997), we opted to characterize spatial-frequency sensitivity with Log-Gaussian shaped tuning functions. Peak-sensitivity was randomly sampled from an exponential distribution – and thus uniformly distributed on logarithmic coordinates – ranging between 1 and 33 c/deg. The average unit bandwidth equalled 1.5 octaves at half height, with a standard deviation of 0.2 octaves. Figure 4-8b illustrates the resulting tuning functions for 15 units from a pool (tuning functions for the other 235 units of the pool have been omitted, for the sake of clarity). The pools in this simulation are thus tuned to a broad range of spatial frequencies. Consequently, only few neurons are optimally sensitive for the 4 c/deg signal, while many are either somewhat sensitive or completely insensitive to the signal.

To test whether our model is able to produce realistic results with a broadly tuned pool consisting of units with correlated noise, 100 pools of 250 units were generated. The spatial-frequency tuning functions, contrast-response functions and inter-unit correlations of these units were determined by randomly selecting parameter values from the appropriate distributions. The effects of noise were approximated as explained in the section on 'Contrast discrimination in noise: effects of response pooling'. Contrast discrimination thresholds in noise are shown for one of these pools in Figure 4-8c. These simulated results are very similar to the findings of Henning and Wichmann (2007).

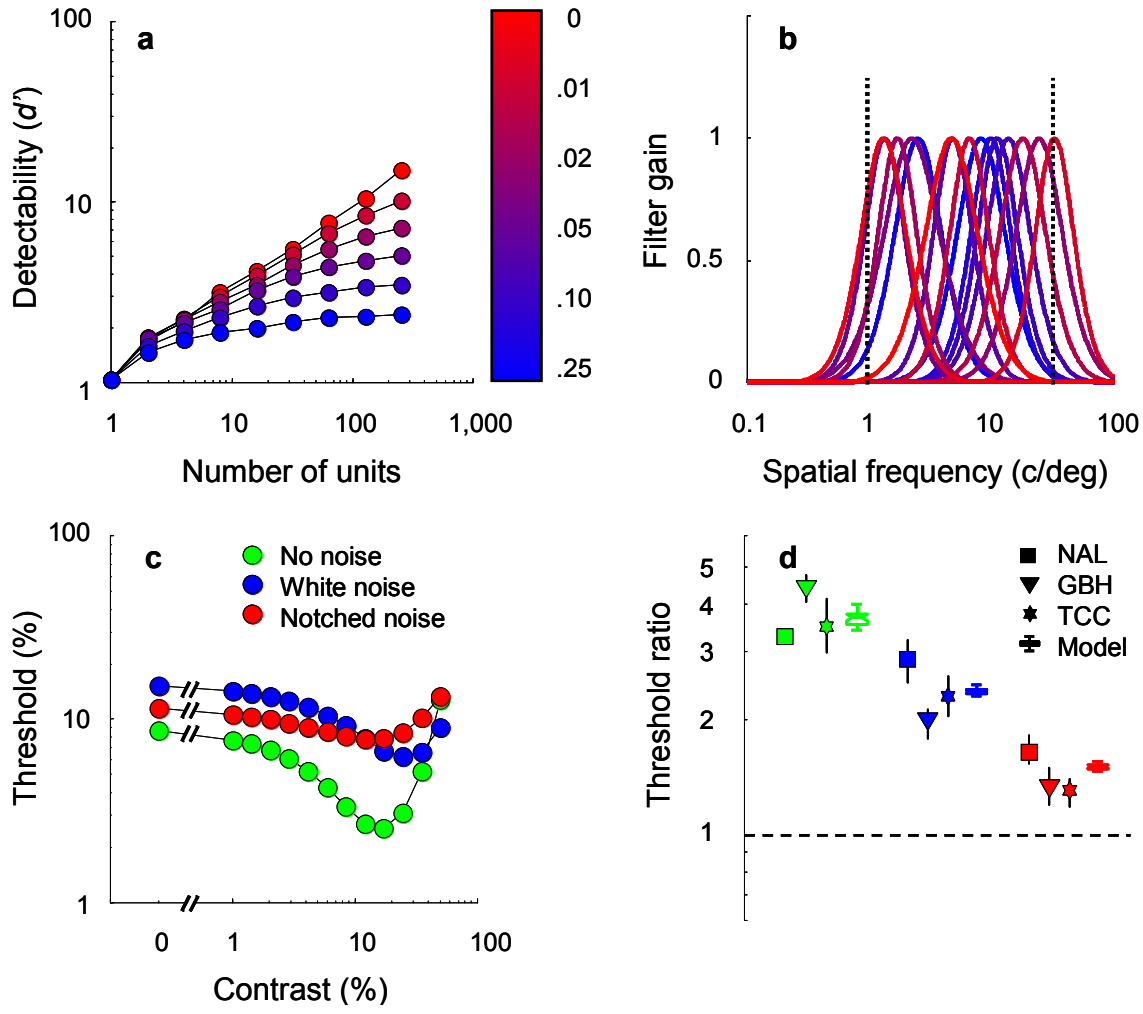

Figure 4-8 *Contrast discrimination in larger pools with random spatial frequency tuning and correlated noise.* ***a.*** *The effect of correlated noise on network performance. Detectability of a low contrast sinusoidal grating is plotted as a function of pool size; colour indicates the mean inter-unit correlation ranging from 0 (the topmost curve in red) to 0.15 (the bottom curve in blue). The standard deviation of this correlation in the simulations was chosen to be one tenth of the average correlation. Each estimate is based on 1,000 pools.* ***b.*** *Spatial frequency tuning functions for 15 units of one particular pool of 250 units (for clarity, tuning functions of the other 235 units have been omitted). The dotted black lines indicate the boundaries of the geometrically uniform distribution from which the unit peak-sensitivities were randomly sampled.* ***c.*** *Simulated contrast discrimination thresholds as a function of pedestal contrast with no external noise (green symbols), in broadband noise (blue symbols) and in notched noise (red symbols) for one particular pool of 250 units.* ***d.*** *75%-correct threshold elevation ratios for the three observers that participated in the Henning & Wichmann (2007) experiments are shown for all noise conditions, indicated by colour. Error bars show the bootstrap-based, non-parameteric estimate of the 68% confidence intervals. Box plots summarize the same factors for 100 pools of 250 units. In all these box plots, the central horizontal line indicates the second quartile (i.e., the median threshold elevation ratio across pools); the other horizontal lines indicate the first and third quartile. Whiskers indicate one and a half times the interquartile range. There were no outliers.*

This can be seen in Figure 4-8d where the threshold elevation – expressing the strength of the pedestal effect – is plotted for all observers in the three noise conditions. The ratio of the 75% correct detection threshold (i.e., no pedestal) to the lowest contrast discrimination threshold (bottom of the dipper) is plotted. Box plots summarise the same ratios in the same noise conditions for the 100 simulated pools. It is clear that our model produces contrast discrimination data that closely mimic the psychophysical data. It can further be derived from this plot that the model behaves very robustly for pools of 250 units, as all box plots are remarkably small, indicating that the different pools produced very similar threshold elevations. We thus conclude that our model robustly produces plausible results for large pools with correlated noise and random spatial frequency tuning.

As explained in the section on weight assignment, response-based weighting produces a dynamical weighting profile, the exact shape of which depends on aspects such as signal contrast, unit selectivity and other factors that affect units' responsiveness. Noise and the spectral characteristics of noise are such factors. It is thus interesting to 'open' the model and compare the weighting profiles in the different noise conditions. Figure 4-9 shows the distribution of weights as a function of the randomly sampled peak-sensitivities for one particular pool of 250 units at a contrast where the signal is barely visible (i.e., the signal contrast corresponds to the bottom of the dipper).

In the absence of external noise (panel a), the weighting profile peaks at the signal frequency (4 c/deg) and is approximately symmetrical on logarithmic coordinates. Units with peak-sensitivities more than one octave away from the signal frequency are usually not very selective for the signal and thus attract almost no weight. With broadband noise added (panel b), the peak of the weighting profile is attenuated, while the tails are elevated because units that are not tuned to the signal frequency are responding to the noise and thus attract some of the weight. With notched noise added (panel c), response-based pooling produces a rippled weighting profile, because the responses of units tuned to frequencies within the notch are most strongly suppressed by the noise. Units tuned to frequencies that are approximately one

octave away from the signal frequency now attract most of the weight. In this regard, notched noise produces more rather than less off-frequency looking in our model – the biggest weights are off the signal frequency but within the notch.

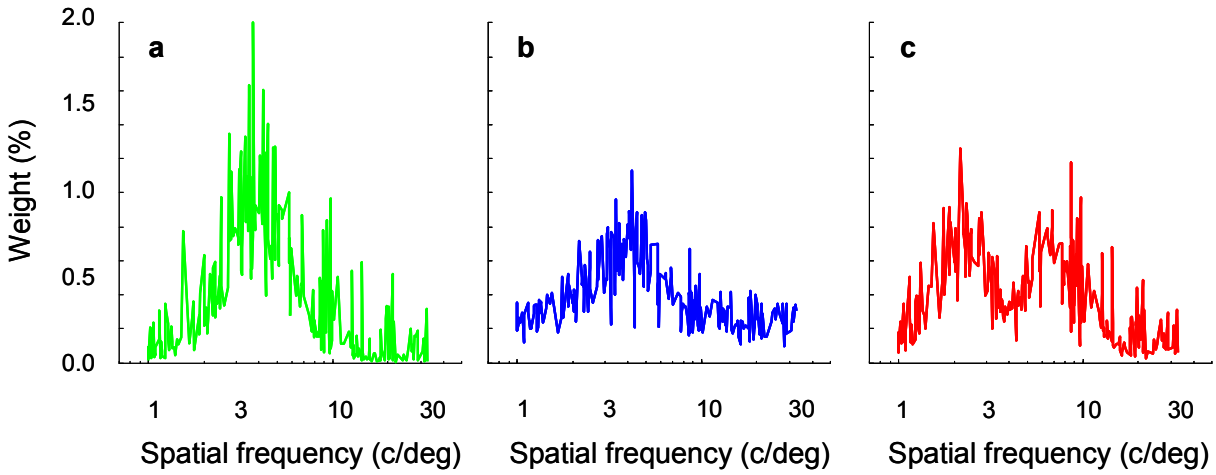

Figure 4-9. *The effect of external noise on the weighting profile. The distribution of weight at a signal contrast corresponding to the bottom of the dipper as a function of peak-sensitivity for one particular pool of 250 units.* ***a.*** *The weighting profile with no external noise added.* ***b.*** *The weighting profile with broadband external noise added to the signal.* ***c.*** *The weighting profile with notched noise added to the signal.*

4.4 Discussion

Current models of spatial vision cannot easily explain why the pedestal effect persists in broadband noise but disappears in notched noise. Indeed, predictions of the standard version of the psychophysical divisive inhibition model (Foley's model 3) are not in line with these findings. Further, there is a discrepancy between the deep dip attributed to single channels by these models, and the mild dip observed in single cells of the striate cortex. If one assumes that notched noise prevents or reduces the use of information carried by channels tuned to frequencies other than the signal frequency, these findings suggest that the pedestal effect stems from off-frequency looking (Henning &Wichmann, 2007). However, the assumption that notched noise prevents off-frequency looking might be wrong. It may be that because divisive

inhibition models were not designed to capture the effects of stimulus variability, the standard version of the gain-control model fails to produce successful predictions.

At single cell level, the normalization model is consistent with many different observations on cortical cell behaviour. Although the available data are limited, this model has also been shown to capture the effects of broadband noise reasonably well. Our implementation of noise effects is fully based on the logic of the normalization model. It is important to note that for notched noise, this implementation does not result in reduced information-pooling across units. To the contrary, notched noise modifies units' contrast response functions, without preventing off-frequency looking. In this case the dipper disappears in notched noise because the notched noise, through the inhibitory effect of noise outside the notch on units tuned to frequencies within the notch, linearizes those units' response and thus removes the pedestal effect.

Here, we have shown that in a network consisting of units whose contrast response functions resemble those of cortical cells, weighted summation – with weights based simply on the magnitude of the response – produces contrast discrimination data that resemble many aspects of psychophysical observations. Similar to earlier neurophysiologically based models of vision, the model includes a spontaneous firing rate (a stimulus-independent base firing rate) and mimics the physiologically-observed proportionality between mean and variance of the firing rate with signal-dependent noise. The pedestal effect in the model, however, arises because of information combination across units.

Using information from non-optimally tuned cells is probably not a unique feature of contrast discrimination. For instance, Shadlen et al. (1996) considered the relation between behavioural and neurophysiological (MT) responses to visual motion and found that non-optimally tuned cells needed to be postulated and included in the neural pool in order to reconcile their behavioural and neurophysiological measures.

Employing a simple response-based weighting heuristic is a sensible strategy, because under most realistic conditions, it resembles an "optimal" combination rule based on weights related to the reciprocal of a covariance matrix. The model predicts not only the standard dipper effect but also how the dipper changes when spectrally flat and notch-filtered noise is added. Finally, the model is consistent with neurophysiological estimates of simple-cell contrast-response functions and thus, irrespective of the specific parameter settings of the model, resolves the puzzling discrepancy between single cell contrast-response functions, which display a weak or even no pedestal effect, and the strong pedestal effect observed psychophysically.

In the model we find that the detailed statistics of the components of the narrowband stimulus hardly matter: performance is principally determined by the signal-to-noise ratio of the decision statistic, which is based on the combined output of both sensitive and relatively insensitive units. Indeed, human contrast discrimination performance has been reported to be largely independent of signal frequency (Bird et al., 2002).

In this paper, we did not discuss effects of orientation tuning. There is, however, no reason to assume that pooling is limited to the spatial frequency dimension. Indeed, the selectivity parameter may be thought off as expressing effects of either or both spatial frequency and orientation tuning. Consequently, our model predicts that performing a contrast discrimination experiment in orientation-filtered noise will produce similar effects as the notched noise effects of Henning and Wichmann (2007).

4.5 Conclusion

Recent evidence has demonstrated that the pedestal effect in spatial vision is differently modified by spectrally flat and notch-filtered noise. Here, we have shown that a network consisting of units whose contrast response functions resemble those of the cortical cells believed to underlie human pattern vision can produce contrast-discrimination data consistent with psychophysical observations when the outputs of multiple units tuned to a range of spatial-frequencies are combined by simple

weighted summation. One implication of these findings is that even in processing low-contrast sinusoidal gratings, the visual system may combine information across neurons tuned to different spatial frequencies and orientations.

4.6 Appendix 4-1

To make the simulations tractable, we simplified the calculation of the weights by using mean responses (Equation 4-3). This simplification allowed us to estimate the variability of the pooled network response (Equation 4-5) and thus network discrimination performance (Equation 4-6) directly for any given combination of pedestal and signal stimulus, without having to simulate too many trials. However, in a real nervous system, weights would necessarily be based on responses alone. Compared to the simplified implementation, there is now trial-to-trial variation in the weights. This additional source of variation does not, however, change important parts of the networks' behaviour described in the paper. The crucial observation here is that in assigning weights, nothing need be known about the variance or indeed any other characteristic of a unit – the weights are determined solely by the strength of the unit's response.

This point is clarified by a simulation, the results of which are shown in Figure 4-A1.1. Figure 4-A1.1a illustrates how the detectability of a narrowband stimulus depends on stimulus contrast. Results for the most selective unit in a pool of 250 units – tuned to a broad range of spatial frequencies and with correlated noise included – are indicated by the purple symbols. Results for the pooled network response as approximated in the paper, i.e., by average-response based weighting, are shown in green. Results for the pooled network response produced by simulating 2,500 trials for all signal contrasts and performing trial-by-trial response-based weighting are shown in black. It is clear that both pooled response functions outperform the most selective unit and are more sharply accelerated.

Consequently, the pooled response functions give rise to a stronger pedestal effect than the single most selective unit whether the effects of the pooling are derived from mean responses or, as would be the case in a real nervous system, from the trial by trial response strength. Estimated contrast discrimination threshold functions are plotted in Figure 4-A1.1b, making use of the same colour conventions as in panel a. Discrimination thresholds were derived from a descriptive function – $d' = \frac{bC^p}{Z^q + C^q}$ –,

fitted to the detectability results shown in Figure 4-A1.1a (the fits are indicated by the coloured lines in panel a). While the thresholds shown in green in panel b lie below the thresholds shown in black at low pedestal contrasts, it is clear that both pooled response functions produce a stronger pedestal effect than the most selective unit. In sum, the simplification in determining the weights is reasonable for the issues discussed in this paper.

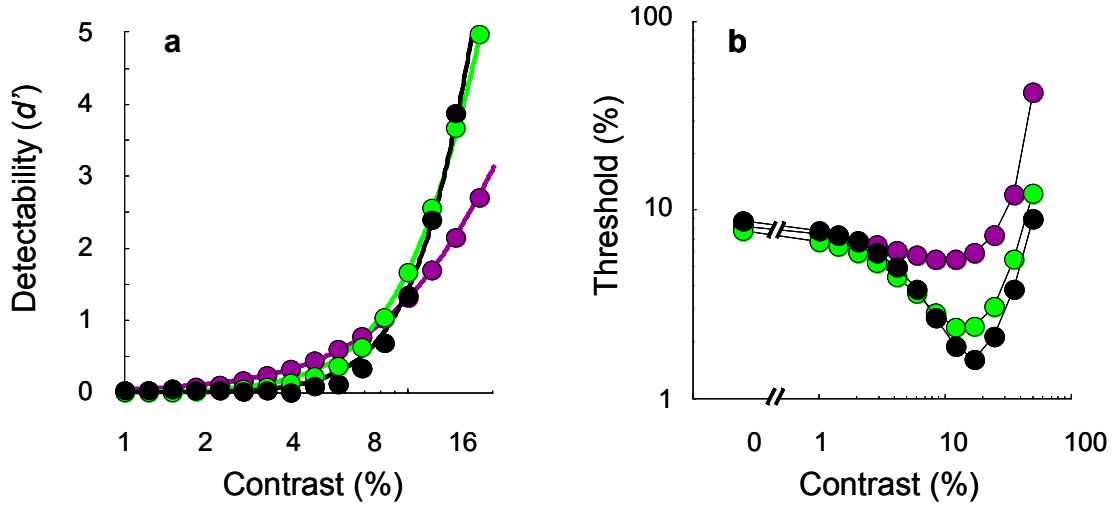

Figure 4-A1.1. *Average-response-based and response-based pooling.* ***a.*** *Detectability, d', expressed as the ratio of the mean response to the standard deviation, as a function of contrast for the most sensitive unit (purple symbols) in a pool of 250 units, average-response-based pooling (green symbols) and trial-by-trial response-based pooling (black symbols) on semi-logarithmic coordinates.* ***b.*** *Contrast discrimination thresholds at 75% correct as a function of pedestal contrast on double logarithmic coordinates for the most sensitive unit (purple), average-response-based pooling (green) and trial-by-trial response-based pooling (black).*

4.7 Appendix 4-2

To capture the effects of external noise on the contrast response function of units in our model, we used theory-based predictions combined with simulations. The basis of most models of cortical neurons is the concept of linear receptive field, followed by an instantaneous nonlinear function. Broadband noise introduces stimulus variability at the preferred spatial frequency and phase of a linear filter and will thus increase the filter's response variance. Because neurons cannot give negative responses, this increased variance also increases the mean cell response at low response rates. At high response rates, this increased variance lowers the mean cell response slightly due to the nonlinearity.

To see all this, consider Figure 4-A2.1, which shows how the contrast response function of a unit to a preferred signal changes with different levels of external noise (panel a shows the mean response, panel b the variance of the response; lighter symbols refer to higher noise levels). To obtain these results, the simulated output of a linear filter was half-wave rectified and passed through the Naka-Rushton equation (see Equation 4-A.1), without any further rescaling:

$$R_{u,bn}(c) = \frac{\left(\max\left(0, R_f(c)\right)\right)^n}{c_{50}^n + \left(\max\left(0, R_f(c)\right)\right)^n}, \qquad (4\text{-A.1})$$

where $R_f(c)$ is the response of a linear filter as a function of signal contrast expressed as a fraction of the maximal response. As in Equation 4-2, the variance of the unit's response was proportional to its mean value.

For the unit shown in Figure 4-A2.1, n equals 2.4 and c_{50} equals 0.38. We performed simulations for a wide range of parameter values and noise levels and found that the results could be captured by resetting the parameters of the unit's response function as given by Equation 4-A.2 (fits to the mean response are shown in Figure 4-A2.1.).

$$\overline{R_{u,bn}}(c) = r_{Noise} + \frac{c^{n-\Delta n}}{\left(c_{50} + \Delta c\right)^{n-\Delta n} + c^{n-\Delta n}}, \qquad (4\text{-A.2})$$

where r_{Noise} is the average noise evoked response and Δn and Δc describe the change of the response exponent and semisaturation contrast in noise. Based on our simulations, these three parameters were estimated analytically for each network unit as a function of external noise level (this is a free parameter in the model), the response exponent n and semisaturation contrast c_{50} (these are randomly selected parameters, as explained in Methods).

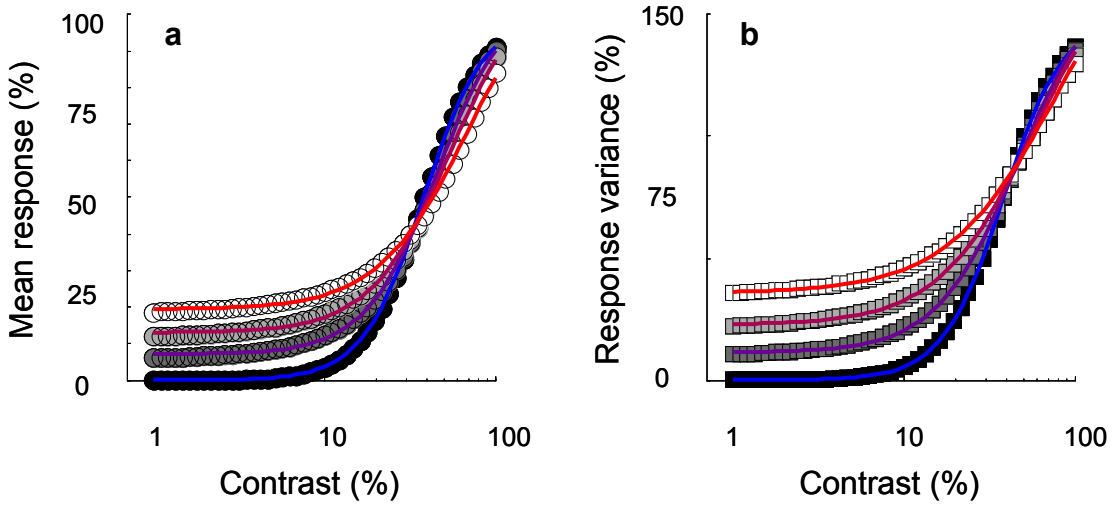

Figure 4-A2.1. ***a.*** *The simulated mean contrast response functions for a given unit with different levels of external noise, making use of Equation 4-A.1 (lighter symbols refer to higher noise levels). Responses are expressed as a fraction of the maximal response. Coloured lines indicate fits of the descriptive model discussed in the text (see Equation 4-A.2).* ***b.*** *The simulated response variance for the same unit.*

Compared to cortical cell data, the noise activation at low firing rates predicted by this simple model is too high. Similarly, the noise suppression at high firing rates is too low (Carandini et al., 1997). Both failures can be cured by an additional parameter α introduced by Carandini et al. to capture non-specific suppression effects of broadband noise, as given by Equation 4-A.3.

$$\overline{R_{u,bn}}(c) = \frac{r_{Noise}}{1+\alpha NC} + Sel\left(\frac{c^{n-\Delta n}}{(c_{50}+\Delta c)^{n-\Delta n} + c^{n-\Delta n} + (\alpha NC)^{n-\Delta n}}\right), \qquad (4\text{-A.3})$$

As in Equation 4-1, *Sel*, expresses the unit's selectivity. Parameter values for α were drawn from an exponential distribution, appropriately scaled to approximate the estimates reported by Carandini et al. At the noise power chosen for the contrast-discrimination simulations in the paper, the response to noise alone was on average approximately four times higher than the spontaneous maintained discharge, r_0. This estimate is reasonably close to the roughly three-fold elevation found by Carandini et al. Further, at this noise level, $\Delta n = 0.45$ (standard deviation of 0.15 across units) and $\Delta c = 0.11$ (standard deviation of 0.03). Figure 4-A2.2 illustrates the effects of broadband noise on the contrast response function for one unit, simulated with the normalization model – i.e., a narrowly tuned excitatory factor and a broadly tuned divisive inhibitory factor, both producing variable responses – and our approximation, making use of Equation 4-A.3.

The effect of notched noise can be inferred from the logic of the normalization model (Heeger, 1997a-b). First, consider a unit tuned to the 1.5 octaves wide notch centred at the signal frequency. Notched noise contains no power at the unit's preferred spatial frequency and phase and will thus not elicit any response by itself. However, the noise has power at frequencies to which the broadband inhibitory gain-control pool is tuned. The normalization model thus predicts that notched noise will produce non-specific inhibition. Given that noise without notch elevates the mean cell firing rate, the inhibitory effect of notched noise is likely to be stronger than the inhibition provided by white noise. For the simulations in the paper, the inhibition provided by notched noise was chosen to be two times higher than the inhibition provided by white noise. (This factor may also partly capture the fact that Henning & Wichmann (2007) increased the power in the pass-bands of the notched noise relative to the white noise for two of three observers).

Because the inhibitory response is largely noise-driven and thus variable, notched noise also modifies the response exponent and semisaturation contrast. As was the case for broadband noise, parameters Δn and Δc used to capture this modification were estimated for each unit based on the simulations of the simple model described above (see Equation 4-A.1). This is of course only an approximation, but sufficient to

capture the increase of Δn and Δc with stronger suppression. This can be seen in Figure 4-A2.2, which illustrates the effects of notched noise on the contrast response function for one unit, simulated with the normalization model and our approximation, making use of Equation 4-A.4.

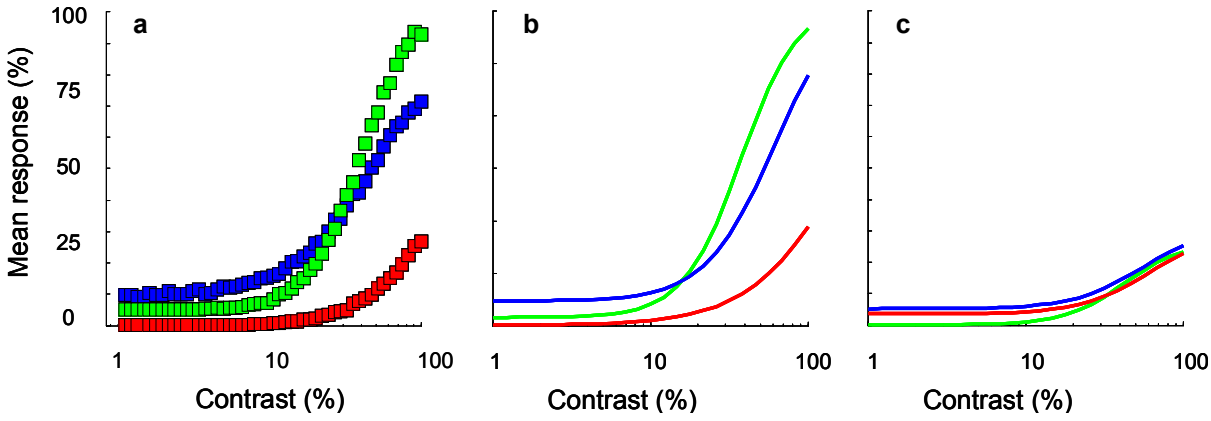

Figure 4-A2.2. ***a.*** *The contrast response functions simulated with the normalization model for a unit tuned to the signal without noise (green), in broadband noise (blue) and notched noise. Without noise, the excitatory and inhibitory factor of the normalization model are constant at any given contrast; in broadband noise, both factors increase and show trial-to-trial variation; in notched noise, the response of the excitory factor is constant, while the inhibitory factor increases and produces a variable response.* ***b.*** *Illustration of the effective mean contrast response functions used in the paper for a given unit without noise (green), in broadband noise (blue) and notched noise (red). This unit is tuned to the signal frequency.* ***c.*** *Same as in panel b for a unit that is not sensitive for the signal frequency. Note that the responses to notched noise and white noise are much more similar for such units.*

The effects of notched noise differ for a unit tuned to one of the pass-bands of the noise (see Figure 4-A2.2, panel c). In short, as the selectivity of the unit for the signal decreases, the notched noise will elicit both excitatory and inhibitory activation and its effects will gradually approximate the effects of broadband noise. The effect of notched noise thus depends strongly on a unit's selectivity. Equation 4-A.4 shows how we implemeted all these effects of notched noise on the contrast response function.

$$\overline{R_{u,nn}}(c) = \frac{(1-Sel)r_{Noise}}{1+\alpha NC} + Sel\left(\frac{c^{n-\Delta n}}{(c_{50}+\Delta c)^{n-\Delta n} + c^{n-\Delta n} + ([1+Sel]\alpha NC)^{n-\Delta n}}\right), \qquad (4\text{-A.4})$$

At the noise level used for the contrast-discrimination simulations in the paper, Δn = 0.43 (standard deviation of 0.3 across units) and Δc = 0.18 (standard deviation of 0.2). The average modification of the response exponent in notched noise thus closely resembles the results in white noise, while the noise-suppression is stronger. For both parameters, the standard deviation in notched noise is higher, due to the effects of tuning. While this implementation captures the main effects of notched noise described above, it is a simplification and at best only an approximation. Nevertheless, this operationalization of notched noise effects proved to be sufficient to generate plausible contrast discrimination data in noise.

For the sake of clarity, the variability in maximal response rate across units was ignored in this appendix. In the model simulations discussed in the paper, however, Equations 4-A.3 and 4-A.4 were multiplied by each unit's r_{max}. Finally, to estimate a unit's response variance, we used the ratio of the variance to the mean simulated in the simple model (Eqation 4-A.1) and multiplied this ratio with the mean responses deduced from Equation 4-A.3 and 4-A.4.

4.8 References

Abbott, L. F., & Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. *Neural Computation, 11,* 91-101.

Albrecht, D. G., Geisler, W. S., Frazor, R. A., & Crane, A. M. (2002). Visual cortex neurons of Monkeys and Cats: Temporal dynamics of the contrast response function. *Journal of Neurophysiology, 88,* 888-913.

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience, 7,* 358-366.

Bird, C. M., Henning, G. B., & Wichmann, F. A. (2002). Contrast discrimination with sinusoidal gratings of different spatial frequency. *Journal of the Optical Society of America A, 19*, 1267-1273.

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

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology (London), 197*, 551-556.

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience, 17(21),* 8621–8644.

Carandini, M., Heeger, D. J., & Senn, W. (2002). A synaptic explanation of suppression in visual cortex. *Journal of Neuroscience, 22(22)*, 10053-10065.

Chen, Y., Geisler, W. S., & Seidemann, E. (2006). Optimal decoding of correlated neural population responses in the primate visual cortex. *Nature Neuroscience, 9(11),* 1412-1420.

Chirimuuta, M., & Tolhurst, D. J. (2005). Does a Bayesian model of V1 contrast coding offer a neurophysiological account of human contrast discrimination? *Vision Research, 45,* 2943-2959.

Deneve, S., Latham, P. E., & Pouget, A. (1999). Reading population codes: a neural implementation of ideal observers. *Nature Neuroscience, 2(8)*, 740-745.

DeValois, R. L., & DeValois, K. K. (1988). *Spatial Vision.* Oxford University Press.

Eliasmith, C., & Anderson, C. H. (2003). *Neural Engineering: Computation, representation, and dynamics in neurobiological systems.* Cambrige: MIT Press.

Finn, I. M., Priebe, N. J., & Ferster, D. (2007). The emergence of contrast-invariant orientation tuning in simple cells of cat visual cortex. *Neuron, 54*, 137-152.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: masking experiments require a new model. *Journal of the Optical Society of America A, 11(6)*, 1710–1719.

Frazor, R. A., & Geisler, W. S. (2006). Local luminance and contrast in natural images. *Vision Research, 46,* 1585-1598.

Freeman, T. C. B., Durand, S., Kiper, D. C., & Carandini, M. (2002). Suppresion without inhibition in visual cortex. *Neuron, 35*, 759-771.

Geisler, W. S., & Albrecht, D. G. (1997). Visual cortex neurons in monkeys and cats: Detection, discrimination, and identification. *Visual Neuroscience, 14(5)*, 897-919.

Golledge, H. D. R., Panzeri, S., Zheng, F., Pola, G., Scannell, J. W. et al. (2003) Correlations, feature-binding and population coding in primary visual cortex. *Neuroreport, 14(7),* 1045-1050.

Goris, R. L. T., Wagemans, J., & Wichmann, F. A. (2008). Modelling contrast discrimination data suggests both the pedestal effect and stochastic resonance to be caused by the same mechanism. *Journal of Vision, 8(15):17, 1-21.*

Goris, R. L. T., Zaenen, P., & Wagemans, J. (2008). Some observations on contrast detection in noise. *Journal of Vision, 8(9):4, 1-15.*

Graham, N. (1989). *Visual Pattern Analyzers*. Oxford University Press.

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: a comparison of single channel and multichannel models. *Vision Research, 11*, 251-259.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: John Wiley & Sons, Inc.

Hammond, P., & MacKay, D. M. (1977). Differential responsiveness of simple and complex cells in cat striate cortex to visual texture. *Experimental Brain Research, 30,* 275-296.

Heeger, D. J. (1992a). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9*, 181-197.

Heeger, D. J. (1992b). Half-squaring in responses of cat striate cells. *Visual Neuroscience, 9*, 427-443.

Henning, G. B., Bird, C. M., & Wichmann, F. A. (2002). Contrast discrimination with pulse trains in pink noise, *Journal of the Optical Society of America A, 19*, 1259-1266.

Henning, G. B., & Wichmann, F. A. (2007). Some observations on the pedestal effect. *Journal of Vision, 7(1:3)*, 1-15.

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

Link, S. W. (1992). *The wave theory of difference and similarity.* Lawrence Erlbaum Associates.

Maffei, L., Morrone, C., Pirchio, M., & Sandini, G. (1979). Responses of visual cortcial cells to periodic and non-periodic stimuli. *Journal of Physiology, 296,* 27-47.

Miller, K. D., & Troyer, T. W. (2002). Neural noise can explain expansive, power-law nonlinearities in neural response functions. *Journal of Neurophysiology, 87,* 653-659.

Montani, F., Kohn, A., Smith, M. A., & Schultz, S. R. (2007). The role of correlations in direction and contrast coding in the primary visual cortex. *Journal of Neuroscience, 27,* 2338-2348.

Nachmias, J. (1981). On the psychometric function for contrast detection. *Vision Research, 21,* 215-223.

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

Olshausen, B. A., & Field, D. J. (2005). How close are we to understanding V1? *Neural Computation, 17,* 1665-1699.

Panzeri, S., Golledge, H. D. R., Zheng, F., Tovée, M. J., & Young, M. P. (2001). Objective assessment of the functional role of spike train correlations using information measures. *Visual Cognition, 8(3-5),* 531-547.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America. A, 2(9)*, 1508-1532.

Pouget, A., Zemel, R. S., & Dayan, P. (2000). Information processing with population codes. *Nature Reviews Neuroscience, 1(2)*, 125-132.

Ringach, D. L., Hawken, M. J., & Shapley, R. (1997). Dynamics of orientation tuning in macaque primary visual cortex. *Nature, 387*, 281–284.

Shadlen, M. N., Britten, K. H., Newsome, W. T., & Movshon, J. A. (1996). A computational analysis of the relationship between neuronal and behavioural responses to visual motion. *The Journal of Neuroscience, 16(4),* 1486-1510.

Shadlen, M. N., & Newsome, W. T. (1998). The variable discharge of cortical neurons: Implications for connectivity, computation and information coding. *Journal of Neuroscience, 18(10),* 3870-3896.

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

Strang, G. (1986). *Introduction to Applied Mathematics.* Massachusetts: Wellesley-Cambridge Press.

Vogels, R., Spileers, W., & Orban, G. (1989). The response variability of striate cortical neurons in the behaving monkey. *Experimental Brain Research, 77,* 432-436.

Wichmann, F. A. (1999). *Some aspects of modelling human spatial vision: Contrast discrimination.* Unpublished doctoral dissertation, The University of Oxford, Oxford, UK.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics, 63*, 1314-1329.

Zohary, E., Shadlen, M. N., & Newsome, W. T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature, 370,* 140-143.

Chapter 5 Some implications for contrast perception models

This chapter addresses some implications of the findings discussed in the previous chapters. First, contrast-detection-in-noise is considered. Detection-in-noise experiments are typically conducted to infer estimates of intrinsic perceptual limitations from models fitted to the data. Here, it is demonstrated that modelling non-linear effects of noise is required to explain contrast detection in noise data at both low and high noise levels. Moreover, modelling these effects forces additional constraints on computational models and modifies certain parameter estimates severely. A more correct characterization of intrinsic perceptual limitations thus requires detection-in-noise experiments and computational models aimed at both low and high noise levels. In the second part of the chapter, implications of the newly proposed population-code model for contrast discrimination are discussed. It is argued that it is not straightforward to implement the standard psychophysical view on channels – having characteristics that do not change during the course of an experiment – in a neurophysiologically plausible environment. Using response-based pooling as a decoding rule may provide a solution because in many cases, predictions of the standard model and this neurophysiologically plausible model may be surprisingly similar.

5.1 Implications for modelling contrast detection in noise

External noise methods and computational models of detection in noise have been used to characterize intrinsic perceptual limitations in a wide variety of applications. Examples vary from perceptual learning (Gold, Bennett & Sekuler, 1999) to developmental dyslexia (Sperling, Lu, Manis & Seidenberg, 2005). Typically, there is a strong focus on the parameter estimates associated with certain conditions. This approach requires the use of correct or at least good-fitting models (Lu & Dosher, 2008). All most prominent contrast-detection-in-noise models predict linear threshold-vs.-noise-contrast functions. The work discussed in Chapter 2 clearly demonstrates noise-induced facilitation for contrast detection in noise and thus rules out purely linear threshold-vs.-noise-contrast functions. Further, the findings presented in Chapter 3 provide additional support for the hypothesis that a single mechanism underlies the pedestal effect and stochastic resonance. Nevertheless, all these results also show that, for detection in noise, the deviation from linear threshold-vs.-noise-contrast functions, especially when considered over a broad noise range, is small. It is thus important to evaluate whether the findings presented in this thesis regarding contrast detection in noise have any implications for external noise methods.

In this section, a quantitative model-based approach is adopted to evaluate whether modelling non-linear effects of noise is required to explain contrast detection in noise data to a reasonable degree. Further, because the main focus of the external noise approach is on parameter estimates, the question whether small changes in values of model parameters suffice to capture the noise benefit is also addressed. To tackle these questions, the data of the three observers that participated in the orientation discrimination task of Chapter 2 are pooled to maximize statistical power. Both a linear and non-linear version of a hybrid nine-parameter contrast perception model are fitted to these data. Model performance is evaluated by using Monte Carlo generated deviance distributions to assess goodness-of-fit. Model performance is further compared in two model selection procedures (AIC and CV). All these analyses will show that characterization of non-linear processes improves quality-of-fit significantly compared to the linear model and is required to describe contrast detection in noise reasonably well.

5.1.1 Data pooling

In experiment 2 of Chapter 2, orientation discrimination performance was measured as a function of external noise level and signal contrast in a 2AFC-task. Figure 5-1a shows how performance changes as a function of signal contrast at the highest noise level for the three observers. Different symbols refer to different observers. The psychometric functions relating the percentage of correct responses to the logarithm of signal contrast were fitted with Weibull functions using the maximum-likelihood procedure of Wichmann and Hill (2001a,b). For the anlaysis in this chapter, the data of the three observers were pooled (see Figure 5-1b).

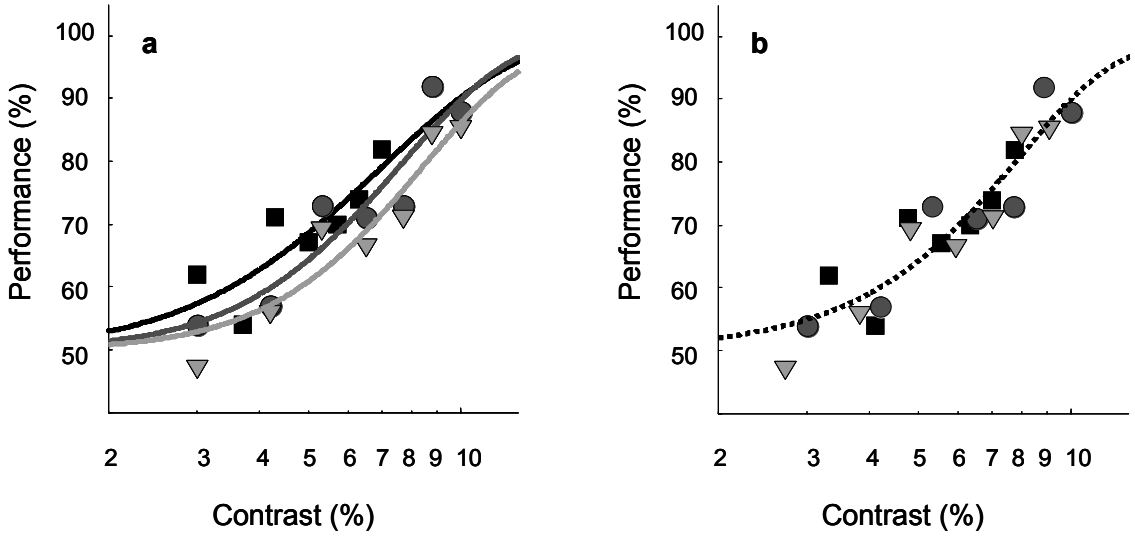

Figure 5-1. ***a.*** *Psychometric functions showing performance (percentage correct) as a function of signal contrast for the highest noise level on semi-logarithmic coordinates. Different symbol shapes and colours are shown for different observers.* ***b.*** *The psychometric function showing performance as a function of rescaled signal contrast for the highest noise level on semi-logarithmic coordinates. To pool the data of all observers, the contrasts of the stimuli were first divided by the 75% correct threshold for each observer and then multiplied by the average threshold of all observers.*

Quality of fit of the psychometric function fitted to the pooled data was assessed for each noise level by comparing the deviance of each psychometric function with a Monte Carlo generated deviance-distribution, as described in Wichmann and Hill (2001a). This statistical test evaluates the null-hypothesis that the data are produced by a stationary observer and thus binomially distributed. First, for the individual

data, quality of fit was always sufficiently high ($p > 0.01$). All observers thus behaved reasonably stable during the experiment. Second, for the pooled data, quality of fit was poor for two noise levels ($p < 0.01$). For these noise levels –the third and sixth, respectively–, the pooled data are thus overdispersed, i.e., these data show more variability than expected from a perfectly stationary observer due to interindividual differences in the slope of the psychometric function. Quality of fit was sufficiently high for the eight other noise levels ($p > 0.01$). Overall, assuming binomial variability for the pooled data-set is thus a reasonable, albeit somewhat optimistic, approximation. This is important, because the contrast perception model used in this paper assumes binomial variance and quality of fit depends on the amount of unexplained variance.

5.1.2 Summary of results

Figure 5-2a shows how orientation discrimination performance changes as a function of signal contrast for three different noise levels: no noise (black symbols), 'optimal' weak noise (grey symbols) and 'moderate' noise (white symbols). Different symbols refer to different observers. The psychometric functions relating the percentage of correct responses to the logarithm of signal contrast were fitted with Weibull functions using the maximum-likelihood procedure of Wichmann and Hill (2001a,b); 10,000 bootstraps were run to estimate the confidence intervals. As can be seen in Figure 5-2b, pooling the data did not affect the shape of the threshold-vs.-noise-contrast functions – plotted for performance levels of 65% correct (blue circles), 75% correct (green circles) and 85% correct (red circles).

We now wish to evaluate whether modelling non-linear effects of noise is required to explain these contrast detection in noise data to a satisfying degree. Further, because the main focus of the external noise approach is on models, we also wish to address the question whether small changes in values of model parameters suffice to capture the noise benefit. To tackle these questions, both a linear and non-linear version of a hybrid nine-parameter contrast perception model were fitted to these data.

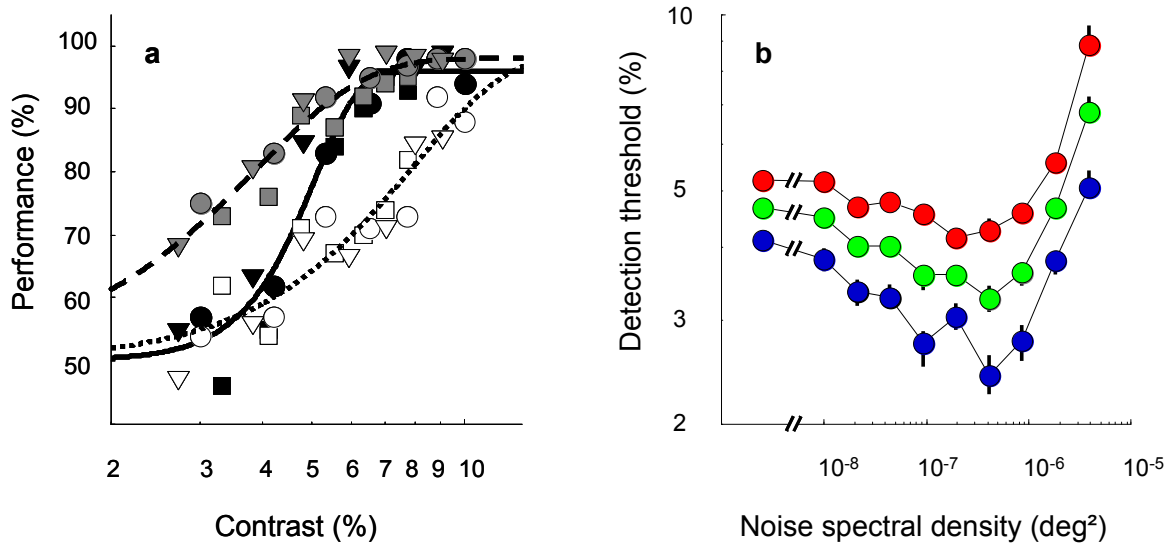

Figure 5-2. ***a.*** *Psychometric functions showing performance (percentage correct) as a function of signal contrast for three different noise levels on semi-logarithmic coordinates: no noise (black symbols), 'optimal' weak noise (grey symbols) and 'moderate' noise (white symbols). Different symbol shapes are shown for different observers. Bootstrap-based deviance distribution estimates reveal that all three psychometric functions fit their respective data well (deviance equals 17.55, 15.56 and 19.82, respectively, which corresponds to the 45th, 29th and 60th percentile of the deviance distribution for a correct model), thus confirming that pooling the data across observers did not introduce an unreasonable amount of variability.* ***b.*** *Threshold-vs.-noise-contrast functions for the pooled data at 65%, 75% and 85% correct (in blue, green and red, respectively) on double logarithmic coordinates. Error bars, when larger than the symbols, indicate ± one SE.*

5.1.3 Two model versions

The model used to fit the data is identical to the model proposed in Chapter 2 (see Figure 2-2 and Section 2.4), except that, as in Burgess and Colborne (1988), early level-dependent noise is included. We will therefore refer to it as a 'hybrid' model (HM). The early noise has both a signal-independent, additive (σ^2_{add}) component and a signal-dependent, induced component with a standard deviation related to that of the external noise. However, in contrast to Burgess and Colborne's findings, an exponential relation turned out to give a much better description of our data than a multiplicative relation, thus $\sigma_{induced} = (\sigma_{ext})^x$. The effective total noise level passed to the sampling and filtering stage is thus: $N_{total} = [\sigma^2_{ext} + \sigma^2_{induced} + \sigma^2_{add}]$. Section 2.4 can be consulted for more information on the model.

Two versions of the HM were fitted to the data with a maximum likelihood procedure (see Section 2.4.2). In the *non-linear HM*, the transducer exponent p was left free to vary. High values of this parameter correspond to a strong expansive nonlinearity and consequently dipper-shaped threshold-vs.-noise-contrast functions (at least for mild levels of early noise). In the *linear HM*, p was constrained to the interval [0, 2], which corresponds to a mild nonlinearity and consequently forced the model to predict linear threshold-vs.-noise-contrast functions and parallel psychometric functions. Because our data require a strong expansive nonlinearity, this operation may be thought off as the effective removal of a free parameter.

While the linear HM is not identical to any of the 'standard' observer models used to characterize contrast detection in noise, it may be thought of as being a representative of this class of models. Indeed, the linear HM produces similar predictions for detection in noise as the standard observer models do. Further, making use of eight free parameters, the model has enough leverage to capture all 'linear' aspects of the data. For instance, the slope of the psychometric function is not determined in advance, as is the case for some of the standard models (Lu & Dosher, 2008). Finally, using two different instances of the same model maximizes model comparison possibilities at the level of parameter estimates.

The fits of both models to the data are shown in detail in the upper row of Figure 5-3 for the three noise conditions also shown in Figure 5-2a. Parameter estimates of both models are given in Table 5-1. In the upper row of Figure 5-3, black lines show the (theory-free) best-fitting Weibull functions and red lines the best fit of the HM. It will be noted that the linear HM overestimates contrast detection performance without noise and underestimates contrast detection performance in the presence of weak noise while the fit to the moderate noise condition is reasonable. The non-linear HM, on the other hand, manages to capture the noise-facilitation effect. At face value, the non-linear HM thus seems to approximate the data much better. This can also be seen in the lower row of Figure 5-3, where summaries of both model fits are shown. Performance, indicated by colour, is plotted as a function of noise spectral density and signal contrast on double logarithmic coordinates. Note that, in the fit of

the non-linear HM, the depth of the dip indeed varies with performance level, while psychometric functions of the linear HM are approximately parallel on semi-logarithmic coordinates. We now wish to assess the quality of fit of each model.

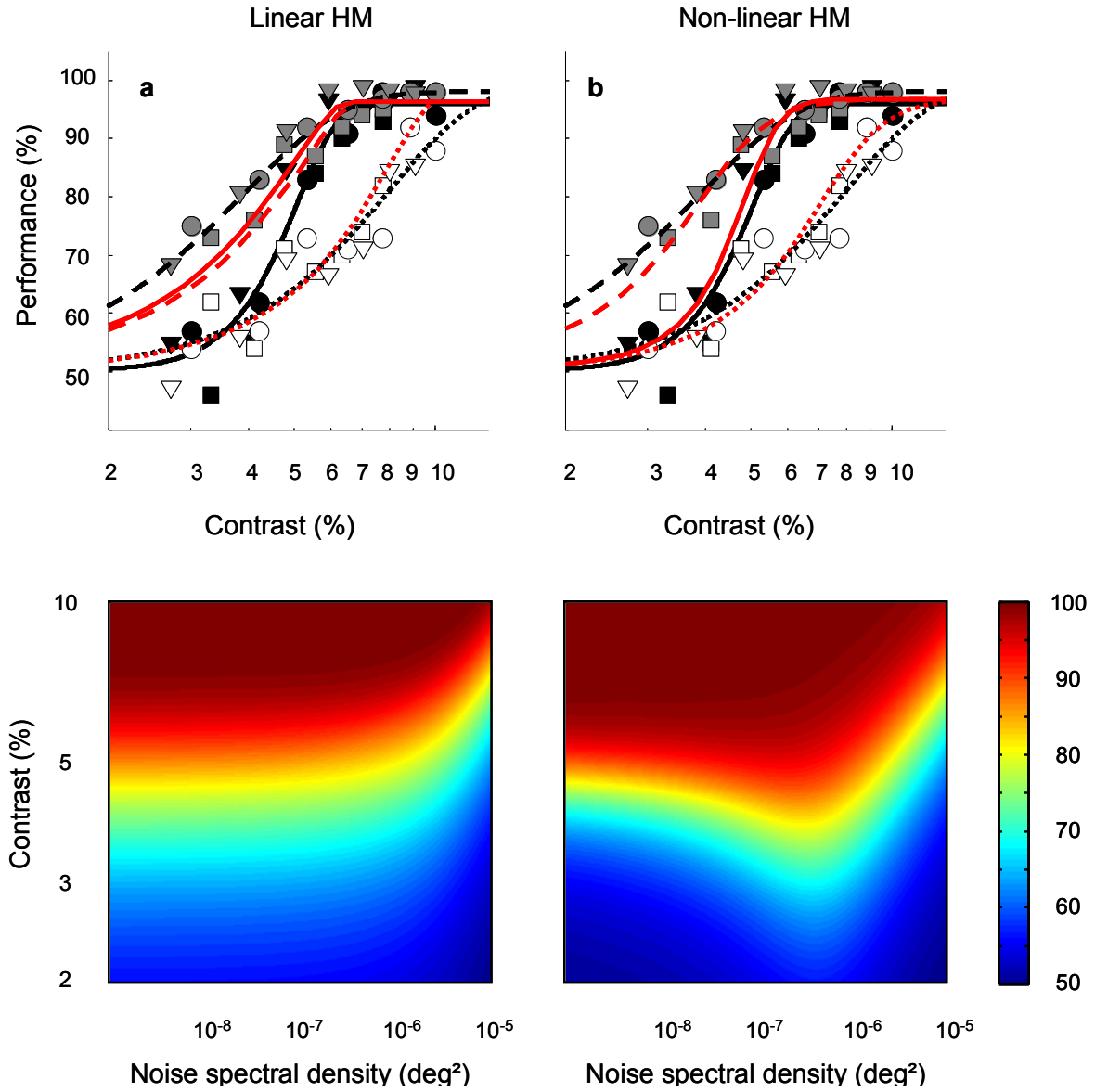

Figure 5-3. *Fits of the linear (left column) and non-linear (right column) HM to the data. The upper row shows detection performance (percentage correct on a linear axis) as a function of signal contrast (percentage Michelson contrast on a logarithmic axis) for the same three noise levels shown in Figure 5-2a: no noise (black symbols), 'optimal' weak noise (grey symbols) and 'moderate' noise (white symbols). Black lines show the Weibull functions fitting the data best, red lines corresponding in style to the black lines, show the model fits. The lower row summarizes the whole model fits. Performance, indicated by colour, is plotted as a function of noise spectral density and signal contrast on double logarithmic coordinates. Note that model predictions differ at weak noise levels, but are virtually identical for noise that is strong enough to produce masking.*

5.1.4 Assessing quality of fit

5.1.4.1 Deviance assessment

The deviance values of both models are shown in Table 1. Wichmann (1999) has shown that, due to the typically relatively small number of measurements, the asymptotically derived deviance distributions often fail to approximate the real distribution for psychophysical data-sets. However, the deviance distribution for small numbers can be estimated easily by means of Monte Carlo simulations. As suggested by Wichmann (1999), we estimated these distributions for both model fits using 10,000 simulated data-sets for an observer whose correct responses in our experiment are binomially distributed as predicted by the model fit. These simulated deviance distributions, as well as the asymptotically derived distributions, are shown in Figure 5-4. Note that the χ^2-distribution with 210 degrees of freedom is too liberal for the linear HM, but too conservative for the non-linear HM, confirming Wichmann (1999) that simulated critical values differ in an unpredictable manner from the theoretical critical values.

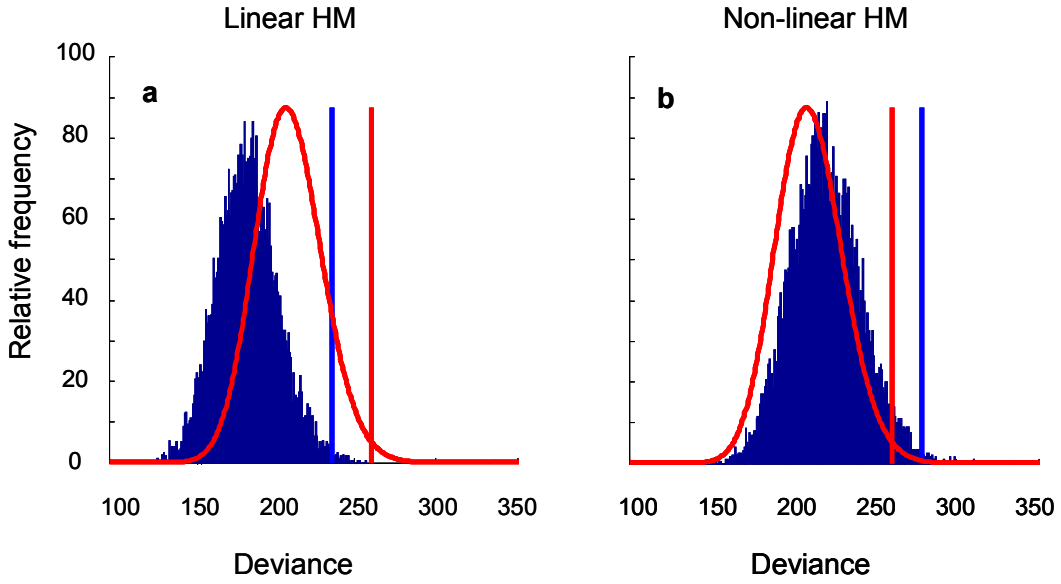

Figure 5-4. *The asymptotically derived deviance distribution (i.e., a χ^2-distribution with 210 degrees-of-freedom, shown in red) and the simulated deviance distributions (the blue histograms, based on 10,000 runs) for the fit of the linear (**a**) and non-linear (**b**) HM. Vertical lines indicate critical values at an alpha-level of 0.01, derived from these distributions. For the linear HM, deviance equals 433 ($p = 0$); for the non-linear HM, deviance equals 276 ($p = 0.016$).*

	x	σ^2_{add}	k	α	p	β	ξ	γ	σ^2_{late}	λ	deviance
Linear HM	1.33	0.004	0.04	4,017	2	0.10	0.01	1.5	**1**	0.04	433.3*
Non-linear HM	0.55	0.026	0.07	3.7e7	7	0.04	3.2e3	1.1	**1**	0.01	275.8

Table 5-1. *The parameter estimates and deviance of the linear and non-linear HM. Bold symbols and numbers indicate frozen parameter values. * indicates a rejected fit at an alpha-level of 0.01. Deviance distributions were estimated for both model fits by means of 10,000 simulated data-sets sets for an observer whose correct responses in our experiment are binomially distributed as specified by the model fit (e.g., Wichmann & Hill, 2001b). Critical values at an alpha-level of 0.01 are 236.5 for the linear HM and 278.8 for the non-linear HM.*

As can be seen in Table 5-1, the fit of the linear model must be rejected at a significance level of 0.01 (p = 0), while the fit of the non-linear HM cannot be rejected at the same significance level (p = 0.016). Thus, despite pooling data across observers –which inevitably introduces some additional variability– the observed deviance for the non-linear HM does not differ significantly from the deviance expected if this model truly underlies the data and the observers' behaviour is perfectly stationary. We now wish to consider the model fits in more detail and assess whether there are any systematic trends in their prediction error (i.e., deviance residuals).

5.1.4.2 Prediction error

In Figure 5-5, deviance residuals are plotted as a function of noise level. The full black line indicates the average prediction error, the dotted lines the 99% confidence interval. This analysis confirms that the linear HM on average overestimates contrast detection performance without noise added and underestimates contrast detection performance in the presence of 'optimal' noise. For the non-linear HM, prediction error does not differ from zero, although there is a trend suggesting that noiseless contrast detection performance may be slightly overestimated. Consequently, the depth of the dip may be underestimated somewhat by the non-linear HM.

Nevertheless, both model evaluation analyses suggest that the non-linear HM gives a reasonable description of our data.

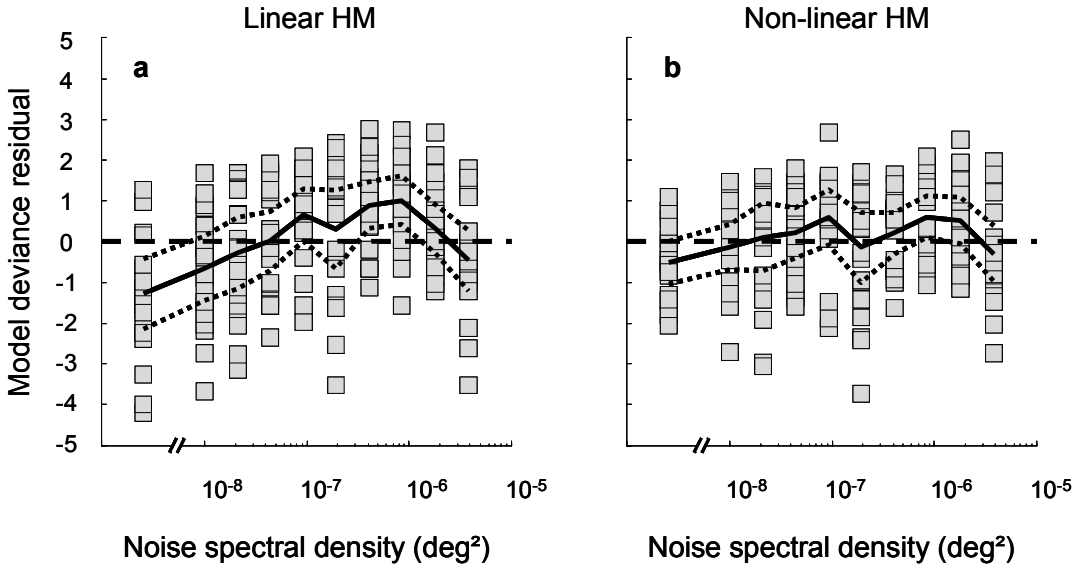

Figure 5-5. *Model deviance residuals as a function of noise spectral density for the linear (**a**) and non-linear (**b**) HM, plotted on semi-logarithmic coordinates. Full lines indicate the mean prediction error, dotted lines the 99% confidence interval. The non-linear deviances (panel b) as a function of the noise spectral density do not differ significantly from zero; the linear model's deviances (panel a) do.*

5.1.5 Model selection

Model selection refers to the problem of selecting, from a group of competing models, the model that best predicts future data, i.e., that generalizes best (see Section 3.4.4). Here, we used *Akaike's Information Criterion* (AIC) and *cross-validation* analysis (CV) as model selection criteria. In the following two analyses, three models were compared: the theory-free Weibull functions (10 noise levels by 3 free parameters, yielding 30 free parameters), the non-linear HM (9 free parameters) and the linear HM (8 free parameters).

Comparison of the Weibull model and the HM should be treated carefully because they do not belong to the same family of models (i.e., they are not nested models). Nevertheless, the Weibull model provides an interesting *upper limit* for model performance, because it consists of the best fitting psychometric function for each

noise level (hence, it can be thought of as a saturated model). Consequently, the model selection analyses will show us whether, at a descriptive level, we gain anything by modelling these data in the first place, i.e., does the HM provide a better quantitative description of our data than a theory-free summary of the data? Second, these analyses will demonstrate whether modelling non-linear effects improves the predictive accuracy of the HM significantly.

For the Weibull model, AIC is 340.1, for the non-linear HM, AIC is 293.8 and for the linear HM, AIC is 449.3. Clearly, the AIC procedure shows the non-linear HM to have the highest predictive accuracy of all three models (despite not being nested models, these AIC-values may be taken to suggest that the Weibull model does over-fit the data). Further, the difference between the linear and non-linear HM is highly significant ($AIC_{lin} - AIC_{non\text{-}lin} = 155.5$; $p \approx 0$). While often used as model selection technique, the AIC measure is only sensitive to one aspect of model complexity, namely number of parameters. Generalizability, however, also depends on other aspects, such as the functional form of a model (e.g., Pitt, Myung & Zhang, 2002).

In (two-fold) CV, the observed data are divided into two subsamples of equal sizes, calibration and validation (sometimes also referred to as training and test set). The calibration subsample is used to estimate the best fitting parameter values of a model. Using these parameter values to calculate the badness-of-fit to the validation subsample yields a model's CV index (e.g., Browne, 2000). The model that minimizes the CV index should be selected. CV is thus an intuitively appealing heuristic method to estimate a model's generalizability. The direct focus on generalizability suggests that CV somehow takes into account the effects of functional form, although it is not clear how it does this exactly (Pitt et al., 2002). Further, large differences in badness-of-fit between the validation and calibration subsample are indicative of over-fitting. And finally, CV enables a fair comparison between models that do not belong to same model family.

We did 1,000 CV-iterations for each model. Each iteration, the 210 blocks of trials were randomly split in two subsamples of 105 blocks. The box plots in Figure 5-6

summarize the normalised deviance of the calibration and validation subsamples for each model. In order to take the number of free parameters used to obtain a given deviance into account, deviance was normalised by *the number of data blocks minus the number of free parameters used to fit the data* (yielding a normalisation term equal to (105 – 8), (105 – 9) and (105 – 30) for the linear HM, non-linear HM and Weibull model in the calibration subsamples, respectively, and 105 for all validation sets). This method compares the deviance to the deviance expected given the number of free parameters used to fit the data (Wichmann, 1999).

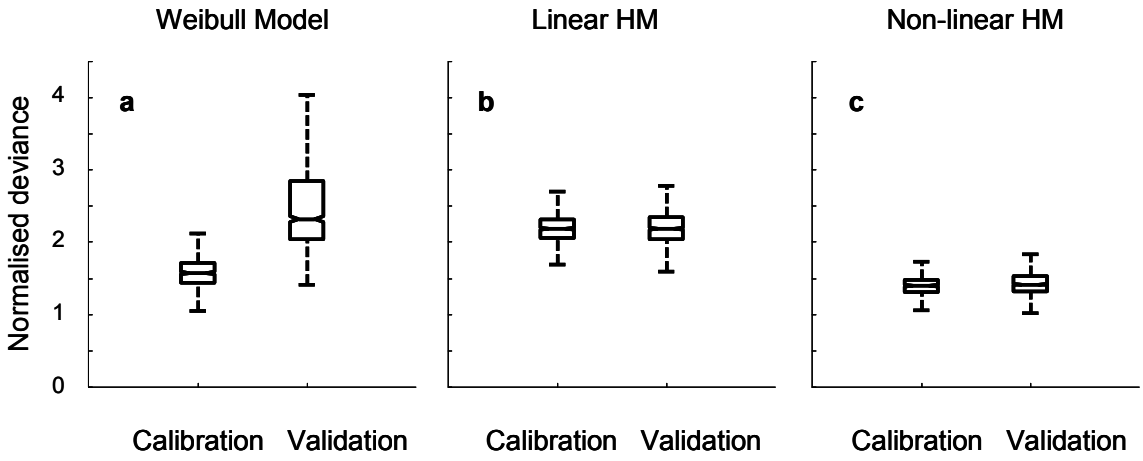

Figure 5-6. *Box plots summarizing the normalised deviance of the calibration and validation subsamples for the Weibull model (**a**), the linear HM (**b**), and the non-linear HM (**c**).*

First, compare the CV index of the three models (i.e., the three rightward box plots). It is clear that the non-linear HM is the best predictor of future data. Second, compare the normalised deviance of the calibration and validation subsample for each model. While largely similar for both instances of the HM, there is a severe difference for the Weibull model. This analysis thus shows that the Weibull model over-fits the data.

In sum, both model selection methods have revealed that, compared to a theory-free descriptive model, the non-linear HM provides a superior quantitative description of our data. Further, modelling non-linear effects of contrast detection in noise improves the predictive accuracy of the HM significantly.

5.1.6 Model comparison

We now consider the parameter estimates and begin with the newly introduced parameter *x*, i.e., the exponent for the early induced noise. In order to compare estimates for both models, it is helpful to 'open' the models and have a look at the early, internal noise. The full lines in Figure 5-7 show how the total level of 2-D, early, internal noise (expressed in standard deviation units) evolves as a function of external noise for both model fits. Estimates for the non-linear HM are plotted in black, for the linear HM in grey. Dotted lines represent an approximation of these functions, making use of a multiplicative instead of an exponential relation between internal induced noise and external noise.

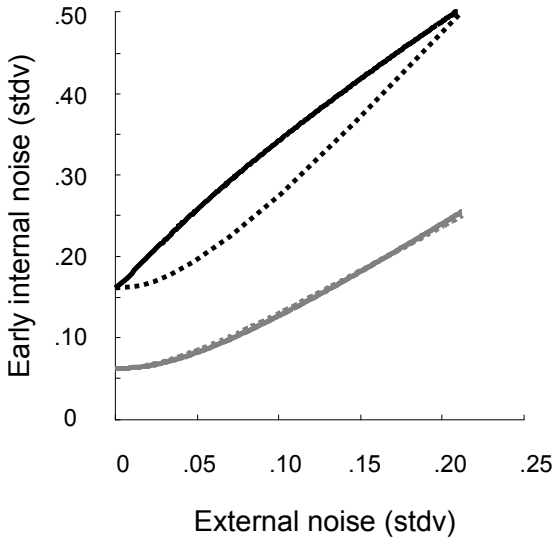

Figure 5-7. *The total level of early, internal noise (i.e.,* $\sqrt{\sigma_{add}^2 + \sigma_{induced}^2}$ *) plotted as a function of external noise level (both expressed in units of standard deviation). The full, black line refers to the magnitude of the early noise that has to be assumed in the non-linear HM to fit the data; the full, grey line shows this assumed magnitude in the linear HM. Dotted lines illustrate approximations of these magnitudes making use of a multiplicative instead of an exponential relation between induced internal noise and external noise.*

First, for the linear HM, multiplicative noise is sufficient to describe the estimated relation between external and early, internal noise, thus confirming Burgess and Colborne (1988) and Lu and Dosher (2008) – note that this is consistent with the

idea that the linear HM indeed is a good representative for some traditional observer models. For the non-linear HM, however, this is not the case. Especially at weak levels of external noise, the shape of the exponential- and multiplicative-noise function differs a lot. Because fits of a multiplicative induced noise model –not discussed further in this paper– were significantly worse than fits of the exponential early noise model ($AIC_{exponential} - AIC_{multiplicative} = 19.8$; $p \approx 0$), the exponential relationship is needed to capture our data. Modelling non-linear effects of external noise thus forces additional constraints on observer models.

Estimates of σ^2_{add}, i.e., the level of early, additive internal noise (when expressed in units of standard deviation) and k, i.e., sampling efficiency, differ by less than a factor three for the models (see Table 5-1). Both parameters have a long history in modelling detection in noise tasks and characterizing observers (Lu & Dosher, 2008; Nagaraja, 1964), although it must be acknowledged that the interpretation of their role differs between models (i.e., in the simplest Linear Amplifier Model –LAM– k and σ^2_{add} capture all visual inefficiencies; in more complex models as the HM they capture only early inefficiencies).

Not surprisingly, the parameters that control the contrast response function (α, β and p) differ considerably between both models. The transducer exponent p has a much higher value in the non-linear application than in the linear version (7 vs. 2, respectively). The positive acceleration in the contrast response function that is directly controlled by p is crucial to fit the noise-induced facilitation effect. However, such strong nonlinear transformation drastically rescales the internal contrast representation dimension. Consequently, the estimates of several other parameters change as well (this is the case for α, a scaling parameter and ξ, the multiplicative factor of the level-dependent late noise).

Estimates of the semi-saturation constant β differ approximately by a factor two for both models. Bear in mind that in the HM, β is expressed in units of *sampled filter response* and not in Michelson contrast – only for an observer who uses a perfect template (i.e., $k = 1$), both scales are identical.

A final interesting result is that the level-dependent component of the late noise is apparently not needed for the linear HM ($\xi = 0.01$), while it plays an important role for the non-linear HM. As will be pointed out in the Discussion, a non-linear model opens the possibility to separate effects of early and late noise.

5.1.7 Discussion

In this section, a quantitative model-based approach was adopted to evaluate whether modelling non-linear effects of noise is required to explain contrast detection in noise data. Given the specific shape of the data and model predictions, the outcome of this approach partly depends on the experimental design: The larger the proportion of measurements gathered at weak noise level – i.e., the noise region where predictions differ for linear and nonlinear models –, the more relevant modelling non-linear processes will turn out to be. Similarly, including only moderate to strong levels of noise would be advantageous for the linear models that capture this part of the data well and are less complex than a non-linear model. However, not measuring effects of weak noise is no convincing evidence for linear threshold-vs.-noise-contrast functions. Consequently, a critical test for the current standard models that predict this linear behaviour thus consists of an experiment specifically aimed at weak to moderate levels of external noise.

The tests performed on the pooled data of experiment 2 clearly demonstrate that goodness-of-fit is acceptable for the non-linear HM, but poor for the linear HM. This lack in quality of fit is due to overestimation of noiseless contrast detection performance and underestimation of contrast detection in weak noise. Two model selection methods (AIC and CV) further revealed that predictive accuracy is significantly higher for the non-linear model than for the linear model. The non-linear model even outperforms a theory-free model consisting of ten psychometric functions – one for each noise-level – in predictive accuracy.

Assuming binomial variance is equivalent to the assumption that rescaling the data removes *all* interindividual differences and this is a somewhat optimistic

characterization of our data-set. One may thus be curious whether the introduction of some additional 'pooling variance' would influence our conclusions. The answer can be found in Figure 5-4. Acknowledging 'pooling variance' would increase the expected deviance for both models and thus shift the deviance distributions to the right. Inspecting the deviance values in Table 5-1 shows that for the nonlinear HM, this operation would increase its plausibility, i.e., the *p*-value would rise somewhat. For the linear HM, however, deviance is so far above the critical value that the *p*-value still would approximate zero.

The linear HM used in this chapter as a representative for the traditional observer models that predict linear threshold behaviour made use of eight free parameters. This may seem fairly complex, given that linear threshold behaviour can already be obtained with two parameters (e.g., Pelli, 1981). Further, inspection of the parameter estimates reveals that not all parameters are used by the linear HM (the two parameters controlling the level-dependent late noise did not affect model predictions in any way and might as well have been left out). One may thus wonder whether the outcome of our modelling exercise would differ for a simpler linear model. The answer is no. The difference in badness-of-fit between the linear and nonlinear HM is simply too big (compare the deviances of both models in Table 5-1). An advantage of the 'complex' eight free-parameter linear HM, however, is that it makes both models extremely comparable: all differences reported are ultimately caused by the difference between the acceleration of both contrast response functions.

We also investigated whether modelling these non-linear effects requires drastic changes in parameter estimates – an issue that is crucial for models. Comparison of both sets of parameter values reveals that the parameters intended to capture visual inefficiencies prior to the non-linearities (i.e., early noise and sampling efficiency) are relatively robust, while parameters that operate after the nonlinearly modified internal contrast scale require severe changes. This makes sense in the light of Birdsall's theorem (Green & Swets, 1966) which states that in a multistage system, sensitive to order, effects of nonlinear transformations occurring *after* the strongest independent noise injection cannot be measured if this noise source is the crucial

performance limiting factor. Thus, once the external noise becomes the crucial performance limiting factor, it effectively removes effects of the nonlinearities and linearizes the system. Consequently, estimates of the component processes influenced by the nonlinearity (e.g., response expansion and late noise addition) will miss crucial factors if they are based only on measurements made in strong noise. Modelling the non-linear effects of external noise, on the other hand, puts additional constraints on the models, as the newly introduced exponential inflation of early induced noise illustrates. Linear modelling of detection in noise already revealed the necessity of induced noise to explain contrast detection performance at high external noise levels (Burgess & Colborne, 1988); but to be able to capture the noise facilitation effect to a satisfying degree, an exponential instead of a multiplicative relation was shown to be necessary.

A second implication of Birdsall's theorem is that a dominant late internal noise source is also a necessary condition for facilitation as a function of noise to occur. Thus, measuring and modelling the beneficial effects of weak, external noise allow one to characterize both early *and* late internal noise, contrary to a purely linear approach (e.g., Lu & Dosher, 2008) or modelling noiseless pattern masking (e.g., Foley, 1994).

5.1.8 Conclusion

Researchers conducting contrast-detection-in-noise experiments typically use computational models to infer perceptual characteristics from their measurements. In this section, both a linear and non-linear version of a hybrid contrast perception model were fitted to the pooled data of three observers that participated in a contrast-detection-in-noise experiment aimed at both weak and moderate noise levels. The hybrid model contains several components of earlier developed spatial vision models. It was first demonstrated that modelling non-linear effects of noise is required to explain contrast detection in noise data at both low and high noise levels. It was then shown that modelling these effects forces additional constraints on computational models and modifies some parameter estimates severely.

5.2 Implications of the population-code model for contrast perception

5.2.1 Ideal observer models cannot explain contrast discrimination…

As discussed briefly in Chapter 1, much of our understanding of near-threshold vision is captured by a model in which a bank of independent, linear spatial-frequency- and orientation-selective filters is followed by a simple maximum-output based decision mechanism – "single-channel models". If no early internal noise is assumed, the Max-rule is equivalent to assuming that the best-tuned channel is monitored (Quick, 1974). This standard view on channels was adopted in the model developed in Chapter 2 and slightly elaborated in Chapter 5 to account for contrast-detection-in-noise. In Chapter 3, it was shown that this kind of model is also able to produce plausible contrast-discrimination-in-noise data for weak and modest levels of spectrally flat noise. A neural implementation of such channel would consist of a group of similarly tuned cells that are each attributed a fixed weight – provided that the contrast response functions of all cells is identical in shape and does not change during the course of an experiment – to produce the channel response.

Henning and Wichmann (2007) demonstrated that the pedestal effect is mildly reduced in white noise, but virtually dissappears in notched noise. These findings raise the important question whether any model attributing sinusoidal contrast discrimination to the characteristics of a single channel provides a description rather than an explanation. Clearly, their results exclude an ideal-observer model in which the spatial weighting profile of the channel corresponds to an optimal signal template: If this were the case, addition of notched noise should not affect contrast discrimination performance at all, contrary to the behavioural results. Thus, channel shapes optimally matched to the spatial frequency and orientation characteristics of the signal – as for instance used in Chapter 2, 3 and 5 – cannot explain contrast detection and discrimination (Henning & Wichmann, 2007). For the model used in these chapters, the implication is that 'sampling' parameter k presumably also reflects the use of a suboptimal filter, for instance a spatial-frequency tuned channel that has an effective bandwidth that is broader than the narrowband Gabor signal or, perhaps, a Mexican-hat shaped weighting function.

5.2.2 …but broadband divisive inhibition models cannot be excluded

However, the claim that the results of Henning and Wichmann (2007) invalidate single-channel models making use of more complex channel shapes, such as broadband divisive inhibition models, may not be correct. This claim is based on assumptions regarding the effects of noise that may not be met in reality. More specifically, Henning and Wichmann assume that notched noise prevents pooling or off-frequency looking while this is not the case for broadband noise.

Many of my own simulations – performed in collaboration with Henning and Wichmann –, none of which is discussed in this thesis, have revealed that there are several difficulties with this explanation of the Henning and Wichmann findings. First and most general, it is not clear how these assumed effects of notched noise can be reconciled with the normalization model of simple cells, or with the divisive inhibition model for single channels. Further, the number of units differentially sensitive to spatial frequency must be limited – otherwise there will be more than one unit tuned to the notch and thus, effectively, pooling in notched noise. Moreover, it is not clear why 1-D noise should prevent pooling across units tuned to different orientations. At a more specific level, my simulations revealed that if the decision statistic is computed by response-based pooling, as proposed in Chapter 4, this explanation requires very specific parameter combinations for the different units in the pool to work. Thus, it is not likely that notched noise prevents off-frequency looking.

Consequently, the claim that these results invalidate *all* single-channel models may be wrong. It cannot be excluded that single channel models can be modified to produce results consistent with the observations made by Henning and Wichmann. For instance, in the model proposed in Chapter 4, for cells tuned to the centre frequency in the notch, notched noise has a stronger linearizing effect than broadband noise. This was derived from the normalization model for simple cells, but could work at the level of a psychophysical channel as well. Thus, while broadband divisive inhibition models are not designed to capture effects of noise, they may be able to explain the Henning and Wichmann findings if modified appropriately. Nevertheless, another approach has been adopted in Chapter 4.

5.2.3 Linking neurophysiology to psychophysics

The motivation for the alternative approach adopted in Chapter 4 is to be found in the superficial relation between the broadband divisive inhibition model for single channels and the normalization model for simple cells: while there is clearly a link between (the arithmetic of) both models, it has never been explored how the neurophysiological and psychophysical levels are related exactly. In the psychophysical version of the model, it is thus not clear how the small dip at the level of single cells is related to the deep dip attributed to psychophysical channels.

This problem was addressed in Chapter 4. It was shown that combining the outputs of many neurons with a small dip produces a deep dip at the level of the decision statistic. Further, due to correlated neural noise, the saturated strength of the pedestal effect closely approximates typical psychophysical findings (see Figure 4-8). In Chapter 4, unit weights are determined by the units' responsiveness. This leads to dynamical spatial-frequency tuning properties at the level of psychophysical 'channels'. Of course, this must not be the case: a weighting scheme arranged to produce static channel tuning may produce similar data – although this was not tested. This type of weighting relates the population-code model to the models used in the other chapters and indeed to the standard view on channels. Thus, one way to connect known neurophysiology to typical psychophysical models is through an appropriate weighting profile.

However, such appropriate weighting profile may be very complex. Obtaining static tuning properties at the psychophysical level will require a dynamic unit weighting profile for the following reasons. First, with fixed weights, interneural variability in the exact shape of the contrast response function (see e.g. Albrecht, Geisler, Frazor & Crane, 2002) will lead to variations in the shape of the channel's tuning function at different contrasts. More specifically, at very low contrasts, poorly tuned neurons may respond louder and contribute more information than better tuned neurons. Second, the contrast response function of simple cells may change severely over relatively small epochs of time, such as one hour – and thus certainly during the course of an experiment. A clear example of this source of variability is shown in

Carandini, Heeger and Movshon's (1997) figure 17. Third, the size of receptive fields of V1 neurons may change depending upon experimental conditions (Das & Gilbert, 1995). Both over short and longer epochs of time, receptive fields in primary visual cortex display many dynamical properties (e.g., David, Vinje & Gallant, 2004; Gilbert & Wiesel, 1992). Thus, implementing a static and contrast-invariant psychophysical channel-shape requires a dynamic and contrast-adjusted weighting profile based on the exact characteristics of each unit. Note further that some behavioural experiments convincingly demonstrate dynamical contrast processing properties at psychophysical level, (e.g., Kwon, Legge, Fang, Cheong & He, 2009).

We opted to use response-based pooling in Chapter 4 because this heuristic combination rule mimics ideal pooling in the absence of external noise and its simplicity is striking – in assigning weights, nothing need be known about any characteristic of a unit. This simplicity strongly contradicts with all the unit characterisctics that need to be known and incorporated in the weighting profile to implement a static, contrast-invariant pscyhophysical channel-shape.

5.2.4 Some implications of response-based weighting as decoding rule

With response-based pooling, the weighting profile – and thus the information passed to the decision statistic – reflects many characteristics of the visual stimulus. If a blank is presented, there is no specific tuning – i.e., the weighting profile is on average flat and reflects only background activity. If a low-contrast signal is presented, there is some mild tuning. If a high-contrast signal is presented, the tuning is more outspoken. If a compound grating is presented, the weighting profile will reflect all components at their respective importance (i.e. contrast) and so on.

This dynamic weighting profile results in a dynamic psychophysical channel-shape, which is at odds with the traditional channel view. Indeed, the success of reponse-based pooling in producing contrast detection- and discrimination data is remarkable and questions the interpretation that should be given to critical band masking studies (e.g., Henning, Hertz & Hinton, 1981; Stromeyer & Julesz, 1972) and other studies

aiming to estimate channel shape (e.g., Blakemore & Campbell, 1969; Legge & Foley, 1980). What do the data gathered in these studies imply exactly? Do they really reflect the 'standard' static channel-shape – as is usually assumed –, or might they also agree with an alternative interpretation in which these results reflect effects of an underlying dynamical channel shape?

Perhaps surpringly, often, data may be in line with both views. At the conceptual level, static channel shapes followed by a Max-rule resulting in a decision statistic reflecting a single-channel response is quite different from response-based pooling resulting in a decision statistic reflecting the responses of many neurons tuned to a variety of spatial frequencies and orientations. Nevertheless, at a qualitative level, many data can be explained with both models. An illustrative example is detection of a compound grating. It is well established that detectability of a grating consisting of two or more components – well seperated in the Fourier domain – that are each just-detectable is remarkably higher than detecability of either of the components (e.g., Graham & Nachmias, 1971; Sachs, Nachmias, & Robson, 1971).

In the single-channel-response view, this increase in detectability is due to probability summation, i.e., the compound stimulus will be detected when any of the – monitored – channels sensitive to one of the individual components detects the stimulus. Because channel noise is assumed to be (largely) uncorrelated, each channel has an independent probability of detecting its component. Adding more components to a grating will then increase its detectability because the probability of *not* detecting the compound is equal to the product of the probabilities of *not* detecting each component.

The response-based pooling decision statistic, on the other hand, also predicts an increase in detectability because adding a second low-contrast component to a single-component grating results in more 'signal' being transmitted to the decision statistic. If the contrast of the second component is sufficiently high to increase the signal-to-noise ratio of some neurons in the pool, the signal-to-noise ratio of the decision statistic will improve and detectability will increase.

Thus, both models may often make similar predictions, at least at a qualitative level. One implication is that experiments that support the single-channel view – and there are very many (Graham, 1989) – typically will not exclude the pooled-responses view. It may thus be an important direction for future work to identify those situations where predictions differ either qualitatively or quantitavely.

As a concluding remark, it should be mentioned that the treatment of both channel views used in this thesis – optimal templates and no predetermined monitoring at all – is probably too extreme. As noted earlier, the first is not consistent with masking effects of notched noise; the second fails to account for results found in uncertainty experiments. For instance, introducing signal-uncertainty on the spatial frequency or spatial position dimension increases detection thresholds somewhat (Davis, Kramer & Graham, 1983). The implication of these findings is that in detection of a known signal, there must be at least some channel monitoring. This is not to say that response-based pooling as a bottom-up weighting scheme is necessarily incomplete. It is reasonable to speculate that channel monitoring originates in a higher-level mechanism, more specifically, attention. Indeed, uncertainty effects have often been interpreted in this way (Graham, 1989).

5.3 References

Albrecht, D. G., Geisler, W. S., Frazor, R. A., & Crane, A. M. (2002). Visual cortex neurons of Monkeys and Cats: Temporal dynamics of the contrast response function. *Journal of Neurophysiology, 88,* 888-913.

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

Browne, M. W. (2000). Cross-validation methods. *Journal of Mathematical Psychology*, *44*,108–132.

Burgess, A. E., & Colborne, B. (1988). Visual signal detection IV. Observer inconsistency. *Journal of the Optical Society of America A, 2*(A), 617-627.

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience, 17(21),* 8621–8644.

Das, A., & Gilbert, C. D. (1995). Receptive field expansion in adult visual cortex is linked to dynamic changes in strength of cortical connections. *Journal of Neurophysiology, 74, 779-792.*

David, S. V., Vinje, W. E., & Gallant, J. L. (2004). Natural stimulus statistics alter the receptive field structure in V1 neurons. *The Journal of Neuroscience, 24(31),* 6991-7006.

Davis, E. T., Kramer, P., & Graham, N. (1983). Uncertainty about spatial frequency, spatial position, or contrast of visual patterns. *Perception & Psychophysics, 33(1)*, 20-28.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: masking experiments require a new model. *Journal of the Optical Society of America A, 11(6)*, 1710–1719.

Gilbert, C. D., & Wiesel, T. N. (1992). Receptive field dynamics in adult primary visual cortex. *Nature 356,* 150-152.

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Signal but not noise changes with perceptual learning. *Nature, 402,* 176-178.

Graham, N. (1989). *Visual Pattern Analyzers*. Oxford University Press.

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: a comparison of single channel and multichannel models. *Vision Research, 11*, 251-259.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics.* New York: John Wiley & Sons, Inc.

Henning, G. B., Hertz, B. G., & Hinton, J. L. (1981). Effects of different hypothetical detection mechanisms on the shape of spatial-frequency filters inferred from masking experiments: I. noise masks. *Journal of the Optical Society of America, 71*, 574-581.

Henning, G. B., & Wichmann, F. A. (2007). Some observations on the pedestal effect. *Journal of Vision, 7(1:3)*, 1-15.

Kwon, M., Legge, G. E., Fang, F., Cheong A. M. Y., & He, S. (2009). Adaptive changes in visual cortex following prolonged contrast reduction. *Journal of Vision, 9(2),* 1-16.

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

Lu, Z.-L., & Dosher, B. A. (2008). Characterizing observers using external noise and observer models: assessing internal representations with external noise. *Psychological Review, 115(1),* 44-82.

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

Pitt, M. A., Myung, I.-J., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review, 109(3)*, 472-491.

Quick, R. F. (1974). A vector-magnitude model for contrast detection. *Kybernetic, 16,* 65-67.

Sachs, M. B., Nachmias, J., & Robson, J. G. (1971). Spatial-frequency channels in human vision. *Journal of the Optical Society of America 61*, 1176-1186.

Sperling, A., Lu, Z.-L., Manis, F. R., & Seidenberg, M. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nature Neuroscience, 8,* 862–863.

Stromeyer, C.F., & Julesz, B. (1972). Spatial frequency masking in vision: Critical bands and spread of masking, *Journal of the Optical Society of America 62*, 1221–1232.

Wichmann, F. A. (1999). *Some aspects of modelling human spatial vision: Contrast discrimination.* Unpublished doctoral dissertation, The University of Oxford, Oxford, UK.

Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness-of-fit. *Perception & Psychophysics, 63*, 1293-1313.

Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics, 63*, 1314-1329.

Chapter 6 General conclusions

The standard view on the early visual system consists of an initial spatial-frequency and orientation-selective filtering stage, followed by a nonlinear post-filter stage. Many behavioural threshold-detection measurements as well as neurophysiological findings support this view. The nonlinear stage is typically attributed to broadband divisive inhibition. To investigate early visual processing in psychophysical experiments, single component sinusoidal gratings are amongst the most popular stimuli. Often, visual noise is added to these stimuli to characterize certain aspects of the visual mechanisms underlying pattern perception. External noise methods are attractive because noise is a random stimulus, yet its average characteristics are highly controllable.

However, effects of visual noise on the processing of sinusoidal gratings are not well understood. At the neurophysiological level, few studies have systematically investigated contrast processing in the presence of noise. At the psychophysical level, assumptions regarding effects of noise vary from excitation to inhibition. Here, visual processing of sinusoidal gratings embedded in visual noise was investigated making use of behavioural experiments and computational modelling.

For detection-in-noise, results of two experiments aimed at weak-to-moderate noise levels revealed that detection thresholds reach a minimum for very low noise levels, i.e., stochastic resonance. At higher noise levels, detection thresholds increase with noise level. These results are consistent with linear-non-linear cascade models such as contrast-gain control models. In these models, the performance improvement arises because rectification causes the external noise to increase the mean filter response to both a weak signal-plus-noise and noise stimulus. In the non-linear post-filter stage, these responses are expanded, which increases the difference between the means of the internal response distributions representing the noise and signal-plus-noise. Because the crucial performance-limiting internal noise source is located in later processing stages, the increased difference between the means improves the signal-to-noise ratio of the decision statistic, resulting in improved contrast detection.

It was shown that a contrast-perception model in which early internal noise is followed by stimulus sampling, linear filtering, non-linear transduction and late noise fits these data reasonably well. This is not the case for models making use of linear contrast transduction. Modelling non-linear effects further proved to force additional constraints on computational models and makes it possible to disentangle effects of early and late noise.

In this elaborated contrast-gain control model, a single mechanism underlies stochastic resonance and the pedestal effect seen in contrast discrimination – i.e., the improved detectability of a sinusoidal grating in the presence of a low-contrast copy of the grating. An alternative hypothesis may be to argue that both effects characterize two independent mechanisms. This issue was addressed by measuring sinusoidal contrast discrimination in weak and moderate noise. When assessed at the 75%-correct detection threshold level, addition of noise did not significantly modify the strength of the pedestal effect.

However, making use of more sophisticated statistical techniques to analyze our data – a full quantitative description of our data with few parameters and comprehensive model selection assessments –, the pedestal effect was shown to be more reduced in weak noise than in moderate noise. This reduction rules out independent, additive sources of performance improvement and supports the hypothesis that a single mechanism underlies both the pedestal effect and stochastic resonance in contrast perception. More specifically, the elaborated contrast-gain control model was shown to produce similar results.

Recently, it has been questioned to what degree models like the contrast-gain control model capture the mechanisms underlying sinusoidal contrast discrimination: these models fail to explain why the pedestal effect persists in broadband noise but almost disappears in notched-noise. While it was argued here that the divisive inhibition model can be modified to capture these effects, it is not clear how the physiological level is related to the psychophysical level in this model. Moreover, implementing a model with static psychophysical channel-tuning properties in a neurophysiologically

plausible, and thus dynamical, environment requires dynamical weighting profiles based on the exact characteristics of each unit.

Here it was demonstrated that the neurophysiological level can be related to the psychophysical level in a much more parsimonious way – yielding dynamical psychophysical channel-tuning properties. In a network consisting of units having the characteristics of the cortical cells believed to underlie pattern perception, a heuristic combination rule that resembles optimal information combination – response-based pooling – produces contrast discrimination data that resemble many aspects of psychophysical observations.

Most notably, this neurophysiologically plausible model explains why the pedestal effect persists in broadband noise, but disappears in notched noise. Moreover, this model reconciles the small pedestal effect observed in single cells of the striate cortex with the deep dip observed in psychophysical channels. These findings suggest that in processing even low-contrast single-component gratings, the visual system may combine information across neurons tuned to different spatial frequencies and orientations.

It was stated in the introductory chapter that the central goal of this thesis is to advance our understanding of the visual processing of sinusoidal gratings embedded in visual noise. All behavioural measurements discussed in the previous chapters are consistent with the following characterization of effects of spectrally flat noise: (1) at low response rates, visual noise increases the mean response; (2) at high response rates, visual noise inhibits the mean response; (3) at all response rates, visual noise increases the response variance.

From a computational point of view, these effects are consistent with the contrast-gain control model. The increased response variance results from the initial linear filtering stage, while the modification of the mean response strength is caused by the post-filter nonlinear stage. Due to the nonlinear operations in the gain-control model, the exact strength of all three effects strongly depends on the parameter values of the

response and gain-control mechanisms and needs to be estimated by means of Monte Carlo simulations. Effects of filtered noise can be derived from the tuning of the excitatory and inhibitory factors in the contrast-gain control model. Different variants of the gain-control model have evolved to become the standard model of cortical cells and psychophysical channels. Nevertheless, the above characterization of noise effects holds for both the neurophysiological and psychophysical level.