UNIVERSITY OF CALIFORNIA, SAN DIEGO

Mechanisms of brightness perception

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Cognitive Science

by

Alan Edward Robinson

Committee in charge:

Professor Virginia de Sa, Chair
Professor Stuart Anstis
Professor Jeff Elman
Professor Don MacLeod
Professor Marty Sereno

2009

The Dissertation of Alan Edward Robinson is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California, San Diego

2009

TABLE OF CONTENTS

Signature Page iii
Table of Contents iv
List of Figures and Tables vi
Acknowledgements vii
Curriculum Vitae ix
Abstract of the Dissertation xi
Chapter 1 1
1. Introduction 1
2. Illusions tested 4
3. The ODOG and UNODOG models 5
3.1. Modeling details 5
3.2. Results—ODOG and UNODOG 5
4. Local normalization of ODOG 8
4.1. Results and discussion 8
5. FLODOG 10
5.1. Results 11
6. Conclusions 13
Chapter 2 16
1. Introduction 16
2. Experiment 1 17
2.1. Methods 17
2.1.1. Subjects 17
2.1.2. Apparatus 17
2.1.3. Stimuli and procedure 18
2.1.4. Results 18
3. Experiment 2 19
3.1. Methods 19
3.1.1. Subjects 19
3.1.2. Stimuli and procedure 19
3.1.3. Results 19
4. Experiment 3 19
4.1. Methods 22
4.1.1. Subjects 22
4.1.2. Stimuli and procedure 22
4.1.3. Results 22
5. Experiment 4 22
5.1. Methods 22
5.1.1. Subjects 22
5.1.2. Stimuli and procedure 22
5.1.3. Results 22
6. General discussion 24
6.1. The speed of brightness processing 24
6.2. Filling-in 26
7. Conclusions 26

Chapter 3 29
Abstract 30
Introduction 31
Experiment 1: Adapt to induction, detect polarity . 38
Methods 39
Subjects 39
Apparatus 39
Stimuli and procedure 40
Analysis 43
Results 44
Experiment 2: Adapt to induction, detect cycle count 45
Methods 46
Results 48
Experiment 3: Jittering edges and slower flicker .. 49
Methods 51
Results 53
General Discussion 54
Conclusions 56

LIST OF FIGURES AND TABLES

Figure 1.1: Illusions Tested 2
Table 1.1: 3
Figure 1.2: ODOG and LODOG models 6
Table 1.2: 7
Figure 1.3: Model Predictions for WE-thick 9
Figure 1.4: The FLODOG model 10
Table 1.3: 11
Figure 1.5: Model Predictions for WE-thick 12
Figure 2.1: Diagram of experimental paradigm 18
Figure 2.2: Experiment 1 19
Figure 2.3: Experiment 1 20
Figure 2.4: Experiment 2 21
Figure 2.5: Experiment 3 23
Figure 2.6: Experiment 4 24
Figure 2.7: Experiment 4 24
Figure 2.8: Diagram of the latency of brightness processing 25
Figure 3.1: Brightness induction and spatial filtering 31
Figure 3.2: Experiment structure 40
Figure 3.3: Aligned and inset conditions 42
Figure 3.4: Experiment 1: induced flicker adaptation 45
Figure 3.5: Experiment 2: count the number of test flickers 48
Figure 3.6: Experiment 3: variable sized adapter 53

ACKNOWLEDGEMENTS

I would like to acknowledge Professor de Sa for her support as the chair of my committee. Her extensive advice has been quite helpful at every stage of this dissertation, from designing the experiments to writing the manuscripts describing the results. It has been a pleasure to collaborate with her.

I would also like to acknowledge the critical role Paul Hammon played in research and development of the FLODOG model presented in chapter 1. The final result is vey much a collaborative effort, and is much stronger thanks to his involvement.

I would like to acknowledge Professor MacLeod for the ideas for the flicker adaptation experiment that led to chapter 3 of this dissertation, and the helpful discussions we have had about the project during its development.

On a more personal level, I would like to thank my parents, Howard and Barbara Robinson, for their encouragement to pursue a career in the sciences, rather than something more practical. I would also like to thank Jymm Gifford for the emotional support provided during the latter half of my PhD.

Finally, I would like to acknowledge that this work was supported by the National Science Foundation under NSF Career Grant No. 0133996 to VR de Sa., and NSF IGERT Grant #DGE-0333451 to GW Cottrell/VR de Sa.

Chapter 1, in full, is a reprint of the material as it appears in Robinson, AE, Hammon, PS, & de Sa, VR. (2007). Explaining brightness illusions using spatial filtering and local response normalization. *Vision Research*, 47, 1631-1644. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in full, is a reprint of the material as it appears in Robinson, AE, & de Sa, VR. (2008). Brief presentations reveal the temporal dynamics of brightness induction and White's illusion. *Vision Research*, 48, 2370-2381. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, has been submitted for publication and is currently under review by the journal Vision Research. The dissertation author was the primary investigator and author of this paper.

CURRICULUM VITAE

Alan Robinson

Doctor of Philosophy, Department of Cognitive Science, UC San Diego, 2009

Bachelors of Arts, Cognitive Science, Hampshire College, 2001.

PUBLICATIONS

1. Robinson, AE, & Triesch, J. (2008). Task-specific modulation of memory for object features in natural scenes. *Advances in Cognitive Psychology*, 4, 1-14.
2. Robinson, AE, Manzi, A., & Triesch, J. (2008). Object perception is selectively slowed by a visually similar working memory load. *Journal of Vision*, 8(16):7, 1-13.
3. Robinson, AE, & de Sa, VR. (2008). Brief presentations reveal the temporal dynamics of brightness induction and White's illusion. *Vision Research*, 48, 2370-2381.
4. Robinson, AE, & de Sa, VR. (2008). Measuring brightness induction during brief stimulus displays. Poster at Vision Sciences Society, Naples, FL.
5. Robinson, AE, & de Sa, VR. (2007). Measuring White's illusion during brief stimulus displays. Poster at Fall Vision Meeting, Berkeley, CA.
6. Robinson, AE, Hammon, PS, & de Sa, VR. (2007). Explaining brightness illusions using spatial filtering and local response normalization. *Vision Research*, 47, 1631-1644.
7. Robinson, AE, Hammon, PS, & de Sa, VR. (2007). A filtering model of brightness perception using Frequency-specific Locally-normalized Oriented Difference-of-Gaussians (FLODOG). Poster at Vision Sciences Society, Sarasota, FL.
8. Robinson, AE, Hammon, PS, & de Sa, VR. (2006). A neurally plausible model of lightness illusions combining spatial filtering and local response normalization. Poster at 13th Joint Symposium on Neural Computation, San Diego, CA.
9. Robinson, A., J. Triesch, M. Hayhoe, J. Droll, & B. Sullivan. (2006). Change blindness during multiple interactions with a single object. Poster at the Vision Sciences Society, Sarasota FL.
10. Robinson, A., A. Manzi, & J. Triesch. (2005). The Costs of Visual Working Memory. Talk presented at the Vision Sciences Society, Sarasota FL.
11. Robinson, A. & J. Triesch. (2004). Visual memory for natural scenes: automatic +

task dependent components. Poster at Vision Sciences Society, Sarasota FL.

12. Robinson, A., & J. Morris. (2002). Is the N170 ERP Component Face Specific? Poster at Psychonomic Society, Kansas City.
13. Robinson, A., & L. Spector. (2002). Using Genetic Programming with Multiple Data Types and Automatic Modularization to Evolve Decentralized and Coordinated Navigation in Multi-Agent Systems. *Late-Breaking Papers of GECCO-2002, the Genetic and Evolutionary Computation Conference*. Published by the International Society for Genetic and Evolutionary Computation.
14. Spector, L., & A. Robinson. (2002). Genetic Programming and Autoconstructive Evolution with the Push Programming Language. *Genetic Programming and Evolvable Machines*, 3, pp. 7-40.
15. Spector, L., & A. Robinson. (2002). Multi-type, Self-adaptive Genetic Programming as an Agent Creation Tool. *Proceedings of the Workshop on Evolutionary Computation for Multi-Agent Systems, ECOMAS-2002*, International Society for Genetic and Evolutionary Computation.
16. Spector, L., R. Moore, & A. Robinson. (2001). Virtual Quidditch: A Challenge Problem for Automatically Programmed Software Agents. In E.D. Goodman, editor, *Late-Breaking Papers of GECCO-2001, the Genetic and Evolutionary Computation Conference*. Published by the International Society for Genetic and Evolutionary Computation.
17. Robinson, A. (2001). Genetic Programming: Theory, Implementation, and the Evolution of Unconstrained Solutions. Hampshire College Division III (senior) thesis.

ABSTRACT OF THE DISSERTATION

Mechanisms of brightness perception

by

Alan Edward Robinson

Doctor of Philosophy in Cognitive Science

University of California, San Diego, 2009

Professor Virginia de Sa, Chair

A physically identical shade of gray on a black background appears lighter than on a white background. This tells us that apparent brightness is not simply a function of how many photons are reflected from a surface, but depends on the surrounding context. This dissertation investigates the mechanisms that underlie this dependence on context.

Chapter 1 presents a computational model of apparent brightness, built out of neurally plausible components. This model uses spatial filtering with oriented difference of Gaussians at several different scales. The output of these spatial filters is locally re-

weighted to normalize the amount of energy within different scales and orientations. This model can account for a wide range of human brightness illusions, using only simple mechanisms. It suggests that brightness perception might be due to relatively early visual areas, and may not require more high-level calculations (such as inferring the 3d structure of the scene), that have been suggested by previous researchers.

If brightness perception is due to early visual areas, then we would expect it to be quite fast. Chapter 2 presents evidence that this is correct. Perceived brightness was measured in human participants who viewed briefly presented stimuli which were then masked to limit the amount of perceptual processing. Subjects were able to report brightness percepts for very brief presentations (as little as 58ms).

If brightness is computed in early visual areas, how is it represented? Chapter 3 asks if brightness is represented in a point-for-point neural map that is filled-in from the response of small, contrast sensitive edge detector cells. Subjects adapted to illusory flicker caused by a dynamic brightness induction stimulus, with a modulating surround and a constant center. Flicker sensitivity was reduced when the test region was the same size as the constant center, but not for smaller, inset regions. This suggests that brightness induction does adapt cells along the contrast edge, but that there is no filled-in population of brightness selective cells to adapt. This is compatible with the model presented in chapter 1, which does not require a filling-in mechanism.

Available online at www.sciencedirect.com
ScienceDirect
Vision Research 47 (2007) 1631–1644
Vision Research
www.elsevier.com/locate/visres

Explaining brightness illusions using spatial filtering and local response normalization

Alan E. Robinson [a,*], Paul S. Hammon [b], Virginia R. de Sa [a]

[a] Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0515, USA
[b] Department of Electrical and Computer Engineering, University of California, San Diego, USA

Received 13 October 2006; received in revised form 5 February 2007

Abstract

We introduce two new low-level computational models of brightness perception that account for a wide range of brightness illusions, including many variations on White's Effect [*Perception*, *8*, 1979, 413]. Our models extend Blakeslee and McCourt's ODOG model [*Vision Research*, *39*, 1999, 4361], which combines multiscale oriented difference-of-Gaussian filters and response normalization. We extend the response normalization to be more neurally plausible by constraining normalization to nearby receptive fields (models 1 and 2) and spatial frequencies (model 2), and show that both of these changes increase the effectiveness of the models at predicting brightness illusions.

Keywords: Brightness; White's effect; Contrast; Computational modeling

1. Introduction

One of the properties that the human visual system extracts is the brightness of surfaces in the visual scene. This is likely an important early stage of visual processing that impacts later stages, such as shape from shading, or even object recognition, where an object must be recognized independently of its illumination. It has been long known that the perceived brightness[1] of a surface depends on the brightness of neighboring surfaces. Fig. 1 shows several examples where identical gray patches appear lighter or darker depending on the immediate surround. These illusions show that the computation of brightness in the visual system requires more than just measuring the amount of light reflected from each surface. Rather, the context, or surrounding surfaces, influences the perceived brightness dramatically. The study of brightness perception, therefore, often concentrates on these kinds of illusions as a way to infer the underlying computational mechanism that drives brightness perception, even when there are no perceptual errors.

There are several different theories of brightness perception. These can be partitioned into high-level and low-level theories. High-level theories suggest that the visual scene is parsed into some kind of meaningful interpretation, and that brightness errors arise as a consequence of how the scene is interpreted. For instance, anchoring theory (Gilchrist et al., 1999), uses perceptual grouping to segment the scene into different visual frameworks, and then scales the perceived shade of gray of surfaces within each framework so that the brightest surface in each appears white, or, at the very least, brighter than without anchoring. Another example of the high-level

* Corresponding author.
E-mail address: robinson@cogsci.ucsd.edu (A.E. Robinson).

[1] Note that the terms *brightness* and *lightness* are sometimes used interchangeably in the literature, though many authors make the distinction that brightness is perceived luminance, and lightness is perceived reflectance (e.g., Gilchrist, 2006). Since the models we consider in this paper do not distinguish between perceived luminance and perceived reflectance, we have elected to just use the term brightness for reasons of simplicity.

0042-6989/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2007.02.017

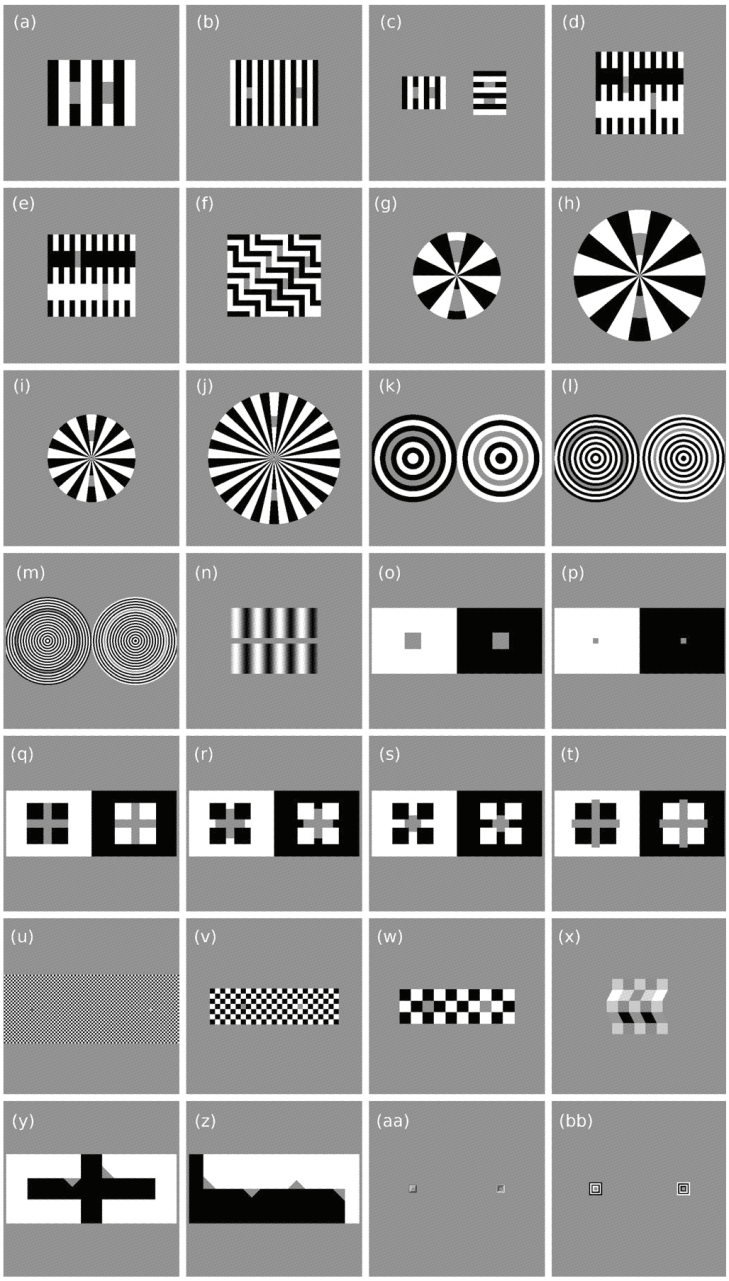


Fig. 1. Illusions tested, all shown to scale, except for (u). Each rectangle is $32 \times 32°$ of visual angle, except for (u) where the scale is twice as large ($16 \times 16°$ of visual angle) and only the central portion of the illusion is shown.

approach is Scission theory (Anderson & Winawer, 2005). According to this theory, surfaces in a scene are split into reflectance, transparency, and illumination layers. The visual system then infers the most probable decomposition into these three layers. Because the decomposition is not always correct, this theory also predicts brightness illusions.

In contrast, low-level theories suggest that brightness errors are caused by interactions of mechanisms in early visual areas which respond to simple features of the image, such as contrast edges, and that no interpretation of the global scene is necessary to cause these errors.

Blakeslee and McCourt (1999, 2001, 2004) and Blakeslee, Pasieka, and McCourt (2005) have introduced and extensively tested a low-level computational model based on filtering by oriented difference-of-Gaussian (ODOG) filters, and then applying global response normalization to equalize the amount of energy at each orientation across the entire visual field. The ODOG model is a compelling starting point because it accounts for many different illusions and uses low-level mechanisms that could be, at least in part, implemented by early visual areas, such as V1. The normalization step in the ODOG model, however, is not particularly neurally plausible, because it is computed globally.

In this work, we extend the ODOG model by exploring whether more neurally plausible normalization schemes will expand the range of illusions predicted by the model. We show that while the normalization step in the ODOG model is necessary to account for a family of illusions known as White's effect (Figs. 1a and b), which are characterized by a highly non-uniform distribution of energy at different orientations, normalization plays relatively little role in predicting non-White's type illusions. Furthermore, we demonstrate that the ODOG model fails on a variation of White's effect that has equal energy at most orientations when integrated across the entire image (Fig. 1c), and also on several previously published variations of White's effect that have a relatively uniform energy distribution across orientation (Figs. 1d–m). These illusions show that the global normalization step in the ODOG model cannot account for all variants of White's effect: instead, more localized normalization schemes may be necessary. A local model of contrast normalization also has the advantage of being more plausible for implementation in early visual areas such as V1.

We introduce two models that add local normalization to the original ODOG model. The first model is locally normalized ODOG (LODOG): instead of normalizing orientation energy across the entire scene, orientation energy is normalized within a local window spanning 4° of visual angle. The second model is frequency-specific locally normalized ODOG (FLODOG): instead of using a fixed window size, normalization is calculated separately for each

Table 1
Sources for illusions tested

Illusion	Fig. 1	Source/original	Test patch size ($w \times h$)	Strength (cd/m^2)
WE-thick	a	Blakeslee and McCourt (1999)/White (1979)	$2° \times 4°$	4.18
WE-thin-wide	b	Blakeslee and McCourt (1999)/White (1979)	$1° \times 2°$	4.6
WE-dual	c	New	$1° \times 2°$	
WE-Anderson	d	Blakeslee et al. (2005)/Anderson (2001)	$1° \times 3°$	6.43
WE-Howe	e	Blakeslee et al. (2005)/Howe (2001)	$1° \times 3°$	0
WE-zigzag	f	Based on Clifford and Spehar (2003)	$1° \times 3°$	
WE-radial-thick-small	g	Based on Anstis (2003)	$2° \times 4°$	
WE-radial-thick	h	Based on Anstis (2003)	$2° \times 4°$	
WE-radial-thin-small	i	Based on Anstis (2003)	$1° \times 2°$	
WE-radial-thin	j	Based on Anstis (2003)	$1° \times 2°$	
WE-circular1	k	Based on Howe (2005)	1° ring width	
WE-circular0.5	l	Based on Howe (2005)	0.5° ring width	
WE-circular0.25	m	Based on Howe (2005)	0.25° ring width	
Grating induction	n	Blakeslee and McCourt (1999)/McCourt (1982)	1° tall	6.23
SBC-large	o	Blakeslee and McCourt (1999)	$3° \times 3°$	11.35
SBC-small	p	Blakeslee and McCourt (1999)	$1° \times 1°$	19.78
Todorovic-equal	q	Blakeslee and McCourt (1999)/Pessoa et al. (1998)	Cross 8° long	2.2
Todorovic-in-large	r	Blakeslee and McCourt (1999)/Todorovic (1997)	Cross 5.3° long	2.4
Todorovic-in-small	s	Blakeslee and McCourt (1999)/Todorovic (1997)	Cross 3° long	4.4
Todorovic-out	t	Blakeslee and McCourt (1999)/Pessoa et al. (1998)	Cross 8.7° long	1.53
Checkerboard-0.16	u	Blakeslee and McCourt (2004)/DeValois and DeValois (1988)	$0.156° \times 0.156°$	7.46
Checkerboard-0.94	v	Blakeslee and McCourt (2004)/DeValois and DeValois (1988)	$0.938° \times 0.938°$	2.84
Checkerboard-2.1	w	Blakeslee and McCourt (2004)/DeValois and DeValois (1988)	$2.09° \times 2.09°$	5.67
Corrugated Mondrian	x	Blakeslee and McCourt (2001)/Adelson (1993)	$\sim 2° \times \sim 2°$	10.85
Benary cross	y	Blakeslee and McCourt (2001)/Benary (1924)	Hypotenuse 3°	9.2
Todorovic Benary 1–2	z	Blakeslee and McCourt (2001)/Todorovic (1997)	Hypotenuse 3°	11.95
Todorovic Benary 3–4	z	Blakeslee and McCourt (2001)/Todorovic (1997)	Hypotenuse 3°	9.55
Bullseye-thin	aa	Bindman and Chubb (2004)	Width 0.608°	
Bullseye-thick	bb	Bindman and Chubb (2004)	Width 0.608°	

Note each illusion is listed in the same order as shown in Fig. 1. Note that the illusion strength listed is the psychophysically measured difference between dark and light patches, averaged across subjects, as reported in the source paper.

frequency and orientation, and the window size depends on the spatial scale of the filter response that is being normalized. Furthermore, each frequency channel is normalized primarily by itself, with decreasing influence from nearby frequencies.

2. Illusions tested

We tested the new models on a wide range of illusions, including many that previous literature has tested with the ODOG model, with varying success. We included examples where the brightness of a test patch is shifted toward the brightness of the region that it shares the majority of its border with, and also examples where the brightness of the test patch is shifted away from the region that it shares the majority of its border with. We will refer to these effects as assimilation and contrast, respectively. Note that we use the terms contrast and assimilation to describe the direction of an illusion, not to indicate the underlying mechanistic cause of the illusion. The underlying causes of these illusions are still of debate.

Except where noted, we duplicated the exact dimensions of each illusion as published in the literature cited, and therefore we will only briefly summarize the relevant details for each illusion. Note that since our goal was to study the ODOG model, we elected to use the illusions as implemented in ODOG-related articles. For this reason, we cite the ODOG-related papers that describe the illusions, as well as the original empirical publication for that type of illusion (Table 1).

To facilitate comparisons between the models and people's perception of the illusions, we summarize here the psychophysical results published in papers by Blakeslee and McCourt. These papers all used a matching paradigm, where target patches on a gray background (Blakeslee & McCourt, 1999, 2001) or a checkerboard background (Blakeslee & McCourt, 2004; Blakeslee et al., 2005) are adjusted by the subjects to match the perceived brightness of test patches in the illusions. While the methods and subjects differ a bit between papers, on the whole the methods are much more similar between these papers than the other sources of illusions we used. This higher degree of methodological similarity allows at least tentative comparisons of the strength of illusions that were tested in different papers, although firmer conclusions can be drawn when comparing data points collected within a single study. In particular, the switch to checkerboard backgrounds around the target patch made some illusions appear as much as 50% stronger than when a gray background was used (Blakeslee & McCourt, 2001). Because such small differences in methodology can have large impacts on the psychophysical results, we elected to not include psychophysical results for illusions published by authors other than Blakeslee and McCourt.

Fig. 1 shows the illusions we tested. The first 13 illusions are all variations on White's effect, and each shall hereafter be referred to as WE-*type*. Except where noted all are seen as assimilation of varying strength. The first two illusions (Figs. 1a and b) are the canonical form of White's effect. Two versions are included because higher frequency versions have been shown to increase the strength of White's effect (Blakeslee & McCourt, 1999). The next 11 illusions are versions of White's effect where the amount of energy at each orientation is more evenly distributed than in the traditional White's illusion. WE-dual (Fig. 1c) is a new configuration of White's effect; the illusion on the right side is just a 90° rotation of the left.

Blakeslee et al. (2005) conducted psychophysical measurements of WE-Anderson (Fig. 1d) and WE-Howe (Fig. 1e). WE-Anderson was found to be weaker than a traditional White's illusion that was exactly matched in terms of test patch size and grating dimensions. With WE-Howe subjects saw either weak contrast or assimilation, with no consistent trend across eight subjects except that people who see White's effects as strong assimilation tend to see WE-Howe as weak assimilation. Note that methodological differences between Blakeslee et al. (2005) and Blakeslee and McCourt (1999) are likely the reason why WE-Anderson appears to be a stronger effect than WE-thin-wide when comparing results between the two papers.

WE-zigzag (Fig. 1f) is based on Clifford and Spehar (2003). This illusion is designed to have nearly equal horizontal and vertical orientation energy locally surrounding the test patches. This is in contrast to WE-dual were orientation energy is only equal when summed over the entire image.

WE-radial (Figs. 1g–j) is based on Anstis (2003). We created several new configurations of WE-radial; the 'thick' (Figs. 1g and h) and 'thin' (Figs. 1i and j) versions are designed to have test patches that are similar to WE-thick and WE-thin, respectively. The 'small' (Figs. 1g–i) and 'large' (Figs. 1h–j) versions denote the radius of the circular grating, which is 8° and 12°, respectively.

Howe (2005) studied a circular version of White's effect where the test patches are embedded in a circular grating shaped like a bull's-eye. The illusion remained when the test patches were extended in length so that they covered an entire ring (making the stimulus similar to that tested by Hong & Shevell, 2004), with almost no reduction in illusion strength. We elected to call this a variant of White's illusion, though the test 'patch' is no longer so analogous to those in a traditional White's effect. WE-circular (Figs. 1k–m) are parametric variations of the illusion, based on the version published in Howe (2005). The illusions are named for the width of the test 'patch' (ring). Subjectively, decreasing the width of the ring appears to increase the strength of the illusion.

We also tested a range of illusions that are not clearly related to White's illusion. We included the Todorovic variations on simultaneous brightness contrast (SBC) (Figs. 1q–t). Blakeslee and McCourt (1999) report that the test patch on the right side of the illusion appears lighter for all configurations except Todorovic-equal (Fig. 1q), where

the patch on the left appears lighter. Note that this does not agree with Pessoa, Baratoff, Neumann, and Todorovic (1998), who report that Todorovic-equal appears lighter on the right, so there is some ambiguity as to the proper prediction for this illusion. For consistency we follow Blakeslee and McCourt (1999).

We also tested several versions of the checkerboard illusion (Figs. 1u–w). The illusion flips between assimilation for Checkerboard-0.16 (Fig. 1u), and contrast at larger spatial scales, which is not captured by Fig. 1 because the figures have been reduced significantly in size relative to laboratory viewing. Note also that Fig. 1u is illegible when shown at the same scale as the other illusions, so in the figure we show just the central portion, enlarged by a factor of two.

For the Todorovic reconfiguration of the Benary cross (Fig. 1z) we list two illusions. This is because the image has four test patches, and our analysis depends on having two test patches per illusion. Thus we split the analysis of this illusion in two, summarizing the results for the two patches on the left and on the right separately. To make clear which test patches we are referring to, we number them 1–4, starting from the left.

The illusions selected here are a representative selection of the illusions that the ODOG model has been tested on previously. In general we elected to include one or two configurations of each illusion, rather than an exhaustive sweep of different scales and relative sizes. Our experience with the models, however, suggests that the results we present will generalize to reasonable variations in the configurations of the illusions.

3. The ODOG and UNODOG models

There are two major stages to the ODOG model. A flowchart of its mechanisms is shown in Fig. 2.

First, the input image is filtered by a set of 42 different filters (Figs. 2a and b). Each filter is a zero-sum difference of Gaussians; the center is circularly symmetric and positive, and the surround is negative and elongated in one direction by twice the extent of the center Gaussian. The filters span six orientations, spaced 30° apart, and seven scales (spatial frequencies), with octave spacing between scales. The largest filters have a central frequency of 6.5 cycles per degree. The filter responses are weighted by the spatial frequency of the filter (in cycles per degree) raised to the power 0.1. This function approximates the human contrast sensitivity function over the frequency range of the filters (Blakeslee & McCourt, 1999). Thus, higher frequency filters receive a higher weight.

In the second stage of the model, the 42 filter responses are summed across spatial scales, generating six different multiscale filter responses, one for each orientation (Fig. 2c). These summed filter responses are then normalized individually by dividing by an image-wide energy estimate calculated as the root mean square (RMS) of the pixels in that summed response (Fig. 2d). This makes the global energy for each orientation equal (Fig. 2e). Finally, the six normalized responses are added together, producing a point-by-point prediction of the relative perceived brightness of the input image (Fig. 2f).

The plausibility of the normalization step is somewhat questionable because, for each orientation, a single normalization factor is calculated over the entire input image. If this computation were to occur in V1, it would require lateral connections or feedback connections that are diffuse enough to allow any part of the visual field to influence responses in any other part of the visual field. It is much more likely that these influences are local, rather than global. Furthermore, in any moderately complex natural image the amount of energy at each orientation is relatively uniform, which would mean that the normalization step would only change the filter responses minimally.

For these reasons, we investigated to what extent the normalization step in the ODOG model is necessary by implementing the model without any normalization, which we call UNODOG (un-normalized ODOG). We then ran ODOG and UNODOG on the set of brightness illusions described in Section 2 to see where normalization played an important role.

3.1. Modeling details

Our implementation of the ODOG model contains two changes from the original Blakeslee and McCourt (1999) implementation. First, when filtering the input image we pad around the edges with gray. Whenever filtering an image, there is always the issue of how to treat the edges; we feel that extending the (gray) background to allow for valid filtering is the most plausible approach. When we tried to replicate Blakeslee and McCourt's exact results we found it necessary to use unpadded convolution, which in effect means the edges are extended by tiling the input. This does not seem particularly likely to occur in V1. In practice, our approach generally led to minor differences, with one exception discussed below.

The other difference is how we calculate the strength of the illusion. Blakeslee and McCourt use the average response along a line cutting through the center of the test patch. Instead, we take the average response for all pixels falling inside the test patch. We elected to use this measure because the values within a test patch are often quite non-uniform, and thus the orientation of the line cutting through the test patch can change the predicted illusion strength. Using all the pixels within the test patch is less arbitrary.

3.2. Results—ODOG and UNODOG

Table 2 shows the predicted illusion strength for the UNODOG and ODOG models (the results for the LODOG model will be discussed in Section 4). To derive a single value representing the predicted strength of each illusion we calculate the difference between the predicted value for the test patch that appears darker and the test

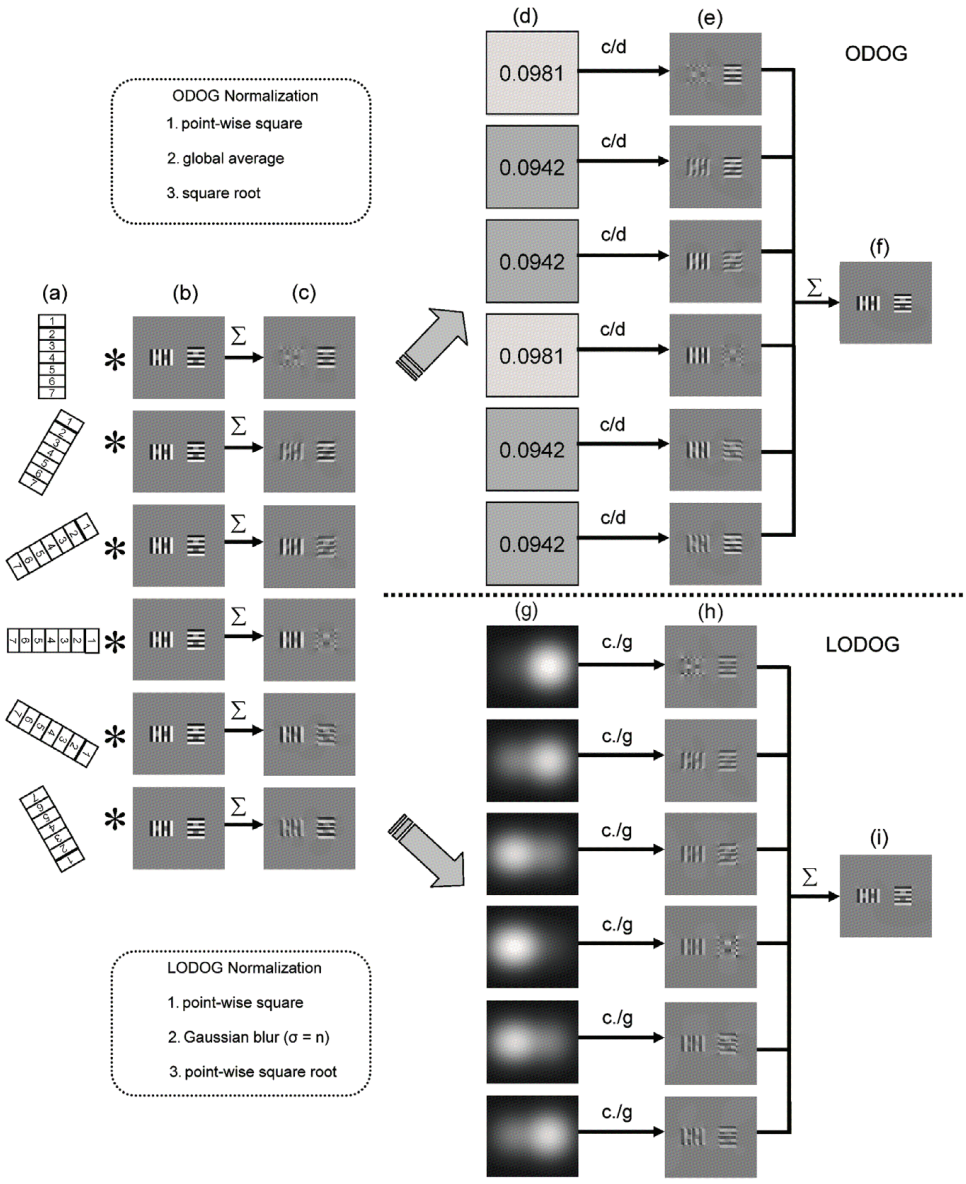


Fig. 2. ODOG and LODOG models. (a) Symbolic representation of the DoG filters at seven different scales and six orientations. (b) The input image. (c) The result of convolving ***a*** and ***b*** and summing the seven scales after weighting the result by a function of spatial frequency. (d) The normalization divisor ODOG calculates for each of the six orientations. (e) The result of applying normalization. (f) Final prediction of ODOG model, produced by summing up ***e***. (g) The point-by-point normalization mask formed by LODOG. (h) The result of dividing each point in ***c*** by each point in ***g*** (./ indicates point-wise division). (i) Final prediction of LODOG model, produced by summing up ***h***.

Table 2
Model results

	ODOG	UNODOG	LODOG $n=1$	LODOG $n=2$	LODOG $n=4$	Illusion Strength (human)
WE-thick	**1.00**	−0.36	**1.00**	**1.00**	**1.00**	1
WE-thin-wide	**2.08**	−0.62	**2.19**	**2.08**	**2.31**	1.1
WE-dual	−0.30	−0.26	**2.53**	**1.36**	**1.11**	
WE-Anderson	−0.15	−1.01	−0.64	−0.30	−0.25	*1.54*
WE-Howe	−0.42	−1.49	−1.99	−0.61	−0.47	*0*
WE-zigzag	−0.51	−0.49	−1.16	−0.76	−0.57	
WE-radial-thick-small	−0.67	−0.67	−0.67	−0.39	−0.55	
WE-radial-thick	−0.41	−0.70	−0.21	**0.01**	−0.29	
WE-radial-thin-small	−0.34	−0.42	**1.43**	**0.21**	−0.20	
WE-radial-thin	−0.22	−0.44	**2.13**	**0.83**	**0.05**	
WE-circular1	−0.82	−1.45	−2.63	−1.04	−1.00	
WE-circular0.5	−0.53	−1.00	−1.47	−0.67	−0.65	
WE-circular0.25	−0.38	−0.71	−1.05	−0.49	−0.48	
Grating induction	**2.03**	**0.17**	**2.32**	**1.69**	**1.77**	1.49
SBC-large	**4.75**	**4.93**	**14.80**	**7.56**	**6.33**	2.72
SBC-small	**6.22**	**6.05**	**26.56**	**14.94**	**9.19**	4.73
Todorovic-equal	−0.36	−0.56	−0.59	−0.26	−0.37	0.53
Todorovic-in-large	**0.49**	**0.77**	**1.63**	**0.55**	**0.52**	0.57
Todorovic-in-small	**0.80**	**1.28**	**2.68**	**0.95**	**0.86**	1.05
Todorovic-out	**0.35**	**0.54**	**1.05**	**0.38**	**0.40**	0.37
Checkerboard-0.16	**1.10**	**0.90**	**2.03**	**0.94**	**0.97**	*1.78*
Checkerboard-0.94	**0.40**	**0.48**	**0.80**	**0.35**	**0.35**	*0.68*
Checkerboard-2.1	**0.69**	**0.72**	**1.62**	**0.60**	**0.59**	*1.36*
Corrugated Mondrian	**0.95**	**0.44**	**2.58**	**0.91**	**0.73**	2.6
Benary cross	**0.09**	−0.12	**0.01**	**0.06**	**0.05**	2.2
Todorovic Benary 1–2	−0.12	−0.41	**0.20**	**0.55**	**0.54**	2.86
Todorovic Benary 3–4	−0.12	−0.41	**0.23**	**0.58**	**0.55**	2.28
Bullseye-thin	−0.74	−0.09	−0.44	−0.35	−0.56	
Bullseye-thick	−0.77	−0.24	−0.52	−0.38	−0.58	

Illusions are listed in the same order as in Fig. 1 and Table 1. Cells in bold indicate the model predicts that the illusion goes in the same direction people typically see it. Note that human values have been scaled so that 1.0 equals the average strength of WE-thick.

patch that appears lighter. We set the sign of the result to indicate whether the prediction matches the direction of the illusion that people see. Negative values indicate that the model predicts the opposite of what people see (i.e., contrast when people see assimilation, or assimilation when people see contrast).

As each model has different normalization steps, the raw numbers that they output are not comparable. To make the results easy to compare between models we scaled the output of each model so that the strength of the WE-thick illusion (Fig. 1a) equals 1, except for the UNODOG model. In this model, WE-thick is predicted in the reverse of what people see, and very weakly. We therefore, instead selected to scale the model's outputs to match the predictions of ODOG on the SBC illusions (Figs. 1o and p). Since the SBC illusion is nearly isotropic, the normalization step in the ODOG model has minimal influence on the strength of the illusion, making these values a good baseline for comparing ODOG to UNODOG.

Since the psychophysics values for these illusions were collected with different methods which impact the strength of the illusions there is no simple way to fairly scale the model output to match the scale of the human responses for all experiments. To enable rough comparison, however, we elected to scale the human data relative to the strength of WE-thick as measured in Blakeslee and McCourt (1999). Keep in mind this decreases how well the models can match to data from the papers after 2001 (shown in italics in Table 2).

The output of the models provides both the predicted direction of the effect (does the test patch get darker or lighter) and also a prediction of the magnitude of the effect. Thus, the models can be judged as to whether they predict the correct direction of the effect, and second, how well the strength of the illusion is predicted. Some care is necessary in using the second metric, as people are highly variable in how strongly they see these brightness illusions. For instance, in Blakeslee and McCourt (2004) data were collected on a White's illusion similar to the dimensions of WE-thin-wide (Fig. 1b). Out of eight subjects, the differences between the two test patches were perceived as small as 2.9 cd/m^2, and as large as 16 cd/m^2. Furthermore, as can be seen in the same paper, the relative strength of different illusions varies somewhat between subjects, even though most of the illusions tested in that paper are White's variants. On the other hand, subjects do tend to see illusions in the same direction, even if the strength varies. Thus, we count the models as being correct if they predict the correct direction of the illusion, although we will discuss the

cases below where the magnitude of the predictions appears to be beyond the range of variability found in human subjects.

By this standard, the ODOG model accounts for the two classic forms of White's effect (Figs. 1a and b). UNODOG, however, does not. It uniformly predicts that people will see contrast instead of assimilation. This shows concretely that the normalization step of ODOG is critical to its prediction of classic White's illusions.

When testing variants of White's illusion with roughly equal global energy at each orientation, we find that ODOG no longer predicts the correct direction of the illusion. These results show that equalizing the orientation energy in the input image makes ODOG fail. Examining the results from UNODOG, we find that it also fails to predict the illusions correctly. Note, however, that the two models do not make identical magnitude predictions, because the different illusion variants have slightly different amounts of energy at different orientations.

The prediction of UNODOG on the grating induction (Fig. 1n) illusion is much smaller than the ODOG prediction, showing that the prediction also depends on unequal energy. Indeed, this reveals that the ODOG model's prediction of grating induction is driven in part by similar mechanisms that make it predict White's effect, and that without normalization ODOG would significantly under-predict grating induction.

Both UNODOG and ODOG predict the correct direction of the SBC illusions (Figs. 1o and p). The models predict, however, that both SBC configurations are about five times stronger than WE-thick (Fig. 1a), when, in fact, the SBC configurations tested here are only slightly stronger than WE-thick (Blakeslee & McCourt, 1999). ODOG clearly over-predicts the strength of contrast, and this is an aspect of the model that needs additional research.

UNODOG and ODOG predict the checkerboard illusions (Figs. 1u and w), with very similar magnitudes. This makes sense since these stimuli are roughly isotropic. Both ODOG and UNODOG account for all of the Todorovic variations of SBC (Figs. 1q–t) except for Todorovic-equal (Fig. 1q), indicating that these predictions do not depend on normalization.

Interestingly, both models predict the corrugated Mondrian (Fig. 1x), though UNODOG predicts it to a smaller extent. This shows that the ODOG account of the Mondrian stimuli depends at least in part on the filters it uses, and not on the complexity of the normalization step. This is a decidedly simple account of an illusion that has been theorized to have high-level origins (Adelson, 1993). Note, however, that both models predict that the Mondrian is weaker than WE-thick, when in fact psychophysical measurements have suggested that it is stronger (comparing psychophysical measurements from Blakeslee & McCourt, 2001 to Blakeslee & McCourt, 1999).

Surprisingly, the Benary cross (Fig. 1y) was predicted correctly by ODOG, but not by UNODOG, revealing that normalization does play a role in predicting this illusion. Neither model predicts the illusion strongly, however, whereas Blakeslee and McCourt (2001) showed that this illusion is not much weaker than the corrugated Mondrian. Interestingly, in contrast to Blakeslee and McCourt (2001), we found that ODOG did not correctly predict the Todorovic version of the Benary cross (Fig. 1z). Upon investigation, we found that this is due to how we padded the input images; if we do not pad the image before filtering (effectively the same as padding the edges with a tiled copy of the illusion) the model makes the correct prediction. Since we find padding with gray to be more plausible, we argue that ODOG does not really account for this illusion. Unsurprisingly, neither does UNODOG.

Finally, ODOG and UNODOG cannot account for the Bullseye illusion (Figs. 1aa and bb), as noted in Bindman and Chubb (2004).

In summary, we found that normalization is key to explaining White's effect, but in general plays a small role in predicting most other illusions considered in this study. For variants of White's effect with more equal global orientation energy, ODOG fails.

4. Local normalization of ODOG

The failures of the UNODOG model show that the normalization step is important for ODOG to account for any of the variants of White's illusion. This normalization, however, is implausible in that normalization of each pixel depends on the energy across the entire scene. We implement a more neurally plausible, local normalization step for ODOG, which we call LODOG (locally normalized ODOG). The mechanisms of LODOG are explained in Fig. 2. The key change is that the summed filter responses are normalized by a local measure of RMS energy instead of a global measure (Fig. 2g). For each pixel, the normalizing RMS is calculated for a Gaussian weighted window centered on that pixel. The window size (n) is specified as the standard deviation of the Gaussian, measured in degrees of visual angle. We tested several different extents to see which sizes of local normalization windows would work as well as global normalization in ODOG.

Local normalization has other advantages as well. Consider the Dual White's illusion (Fig. 1c). While the illusion strength appears to be undiminished, the global energy is now nearly equal for each orientation, so the global normalization step of ODOG will have little effect. Since normalization is key to ODOG predicting White's illusion, this means that ODOG fails to make the correct prediction. Since LODOG uses a local window, each copy of White's illusion in Fig. 1c will be normalized relatively independently.

4.1. Results and discussion

We tested the LODOG model with Gaussian normalization windows of standard deviation $n = 1$, 2, or 4 ° of

visual angle. The predictions of the model are shown in Table 2. To facilitate comparison across models, the model outputs are scaled so that the illusion strength of WE-thick (Fig. 1a) equals 1.0.

For all the White's family of illusions that ODOG correctly predicts, LODOG also predicts that the illusions go in the same direction, independently of the window size tested. In contrast to ODOG, LODOG also can predict WE-dual in the correct direction, with the smaller window sizes predicting stronger illusions. LODOG does not, however, predict some of the equal energy variants like WE-Anderson (Fig. 1d), WE-Howe (Fig. 1e), or WE-zigzag (Fig. 1f). LODOG's performance is somewhat better on the WE-radial illusions (Figs. 1g–j), where some of the stimulus configurations are correctly predicted by some of the window sizes, but not all. LODOG also did not improve performance on WE-circular (Figs. 1k–m).

To summarize, LODOG does fix the simple case WE-dual, but for more complex equal-energy White's variants, its performance is only mixed, though it never does worse than ODOG.

LODOG also predicts SBC (Figs. 1o and p), but overestimates the illusion's strength, especially for small normalization windows. ODOG also over-predicts the strength of SBC, and the prediction for LODOG with a window of 4° is not much worse than ODOG predictions. Clearly, local normalization does not improve the ability to predict the strength of SBC. LODOG predicts the checkerboard illusion (Figs. 1u–w), with stronger predictions made when the window size is smaller. LODOG predicts the Todorovic variants of SBC (Figs. 1q–t) about as well as ODOG does.

LODOG predicts a smaller effect for the Benary cross (Fig. 1y) than does ODOG. For the Todorovic variation of the Benary cross (Fig. 1z), however, LODOG (with window sizes of 2° or 4°) predicts the illusion in the correct direction, something that ODOG, as we implemented it, does not. Thus, LODOG appears to be somewhat better than ODOG at predicting the Benary cross across variations in configuration, but is not a complete explanation of the effect, since it predicts a fairly small illusion.

Finally, LODOG's predictions on the Bullseye stimuli (Figs. 1aa and bb) are no better than ODOG's.

In addition to the mean illusion strength predicted by the LODOG models, we also examined the point-by-point predictions made by LODOG with differing window sizes. Fig. 3 shows the cross-section of the model's predictions for WE-thick. The cross-sections show that as the size of the normalization window decreases, the predicted uniformity of regions is decreased, relative to ODOG, and that smaller window sizes affect cross-sections by increasing the depth of valleys and the sharpness of peaks, with this effect becoming very extreme for ($n = 1°$). In Blakeslee and McCourt (1999), psychophysical data were collected for the test patches, and it was found that there was a non-uniform gradient of brightness across the test patch, which was similar to what ODOG predicted. In fact, the

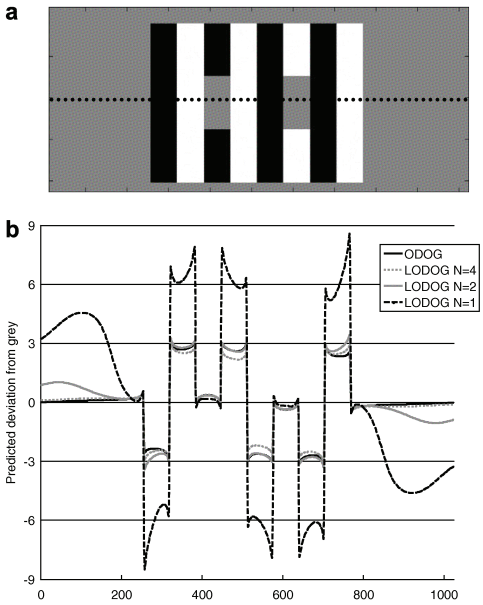


Fig. 3. Model predictions for WE-Thick. (a) Dotted line indicates location of cross-sections. (b) Perceived brightness predicted by ODOG and LODOG models along the cross-section.

psychophysical values suggested that the brightness profile had slightly more curved (i.e. deeper) valleys than ODOG predicted. Thus, the predictions of LODOG with a large window ($n = 4°$) may actually be closer to the psychophysical values than ODOG's.

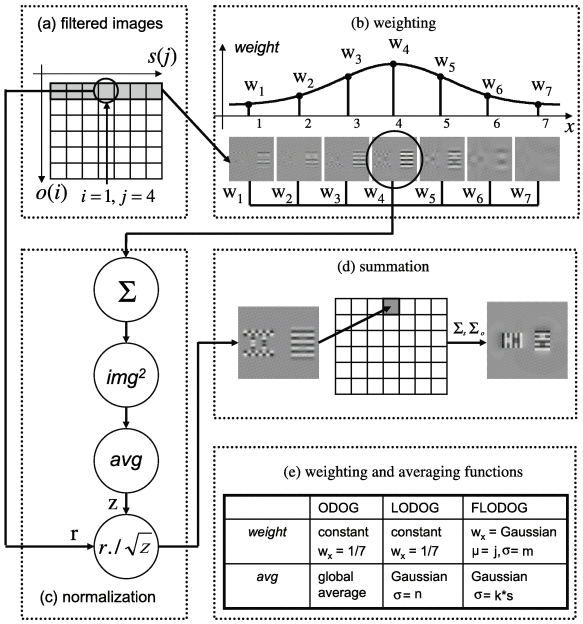
Taken together, these results show that LODOG, in general, works at least as well as ODOG, especially when the window size is larger, such as 4°. This shows that the ODOG normalization step can be made local without reducing the ability of the model to predict a range of brightness illusions. In fact, this more plausible localization scheme actually allows LODOG to predict some illusions that ODOG does not. There remain, however, many equal energy variants of White's illusion that LODOG does not account for. If a normalization-based model is to account for these illusions, a more complex extension to ODOG is necessary.

Examining the results we hypothesized that there were two extensions which would make the model more biologically plausible and which might also improve its ability to predict the illusions. One is that the size of the normalization window should not be constant; rather, very small-scale filters should be normalized by smaller local regions than large-scale filters. Second, it has been shown that neurons with similar spatial frequency preferences tend to cluster together in V1 (Issa, Trepel, & Stryker, 2000; Tootell,

Silverman, & De Valois, 1981; see also Sullivan & de Sa, 2003). Thus, lateral inhibition between neurons should be biased toward neurons of the same spatial frequency. This suggests a new way to normalize the response of a filter, which is local in spatial terms, and also localized to nearby frequencies. We implemented a new model called FLODOG (frequency-specific locally normalized ODOG), which implements these two changes. While all of these changes are reasonable, and likely to make the model more neurally plausible, they do increase the complexity of the model. In the next section, we will evaluate how well this new model compares with the simpler LODOG model.

5. FLODOG

The FLODOG model extends the LODOG model by adding frequency-dependent normalization windows and local weighting when summing across scales. Fig. 4 outlines how FLODOG works, and how it relates to the ODOG and LODOG models. The differences between the FLODOG and ODOG models start after the 42 filter responses have been generated and weighted by spatial frequency (Fig. 4a). For each filter response (r) a new normalization mask is created. Instead of summing across all scales for a single orientation, a weighted sum across frequencies is used. Each filter weight w is computed using a Gaussian function with standard deviation m, shifted so the highest point is centered on the filter being normalized (Fig. 4b). The sum of the weights is normalized to 1. The weighted sum is converted to a localized energy estimate by squaring the value at each point, blurring the whole image by a Gaussian of standard deviation n, and then taking the square root, point-by-point. n is calculated for each filter response by multiplying s (the standard deviation of the center of the DoG filter that generated r) by a scalar, k. This process generates a point-by-point local energy estimate for each r that also includes energy from nearby scales. r is normalized by dividing each point by square root of z, the local energy estimate for that point (Fig. 4c). Finally, each normalized filter response is summed up to produce a point-by-point estimate of the perceived brightness (Fig. 4d). Though it would be computationally inefficient, ODOG and LODOG can be thought of as operating in the same way as FLODOG does, except the weight w would be constant, and different averaging equations would be used (Fig. 4e). The averaging equation for ODOG is the global root mean square of the summed filters, whereas LODOG uses the same averaging equation as FLODOG except that n (the standard deviation of the blur) is a constant, independent of spatial scale.



	ODOG	LODOG	FLODOG
weight	constant w_x = 1/7	constant w_x = 1/7	w_x = Gaussian μ = j, σ = m
avg	global average	Gaussian σ = n	Gaussian σ = k*s

Fig. 4. The FLODOG model. (a) The 42 filter responses, which are generated identically to ODOG. (b) Example of weighted average calculated for a single filter response, at orientation $i = 1$ and scale $j = 4$. Note that the shape of the Gaussian changes for other values of j. (c) Normalization is applied to a single filter response, using the weighted average from (b) to calculate the local energy. (d) All 42 normalized filter responses are summed to produce the final prediction. (e) A comparison of the different weighting and averaging functions used by ODOG, LODOG, and FLODOG.

We tested a variety of FLODOG parameter combinations, crossing the size of the normalization window ($n = 2s$, $3s$, or $4s$) with the weighted sum across frequencies ($m = 0.25$, 0.5, 1, 1.5, 2, and 3). We tested these models on the same illusions we used for ODOG and LODOG.

5.1. Results

We found that FLODOG performed well across a wide range of parameters. The scaling of the normalization window between $n = 2s$ and $n = 4s$ had minimal effect on the predictions of the model for many of the illusions, with the notable exception of SBC and Bullseye. In contrast, the weighting of nearby frequencies (m) makes a big difference in which illusions are predicted correctly and also the magnitude of the predictions.

Due to space considerations we present a subset of models that we tested (Table 3). We include the model that accounted for the most illusions, FLODOG with a normalization window of $n = 4s$, and a weighting of nearby frequencies of $m = 0.5$. For comparison we also include FLODOG with $n = 2s$, $m = 0.5$, and $n = 4s$, $m = 3$. To allow comparison between different models we also include the response of ODOG, and of the most successful LODOG model (window size of 4°).

FLODOG with $m = 0.5$ accounts for all the White's illusions that LODOG does. In addition, it predicts the correct direction of illusion for WE-zigzag (Fig. 1f) and all of the WE-radial (Figs. 1g–j) and WE-circular (Figs. 1k–m) illusions. None of the models we tested, however, predicted assimilation for the Anderson (Fig. 1d) or Howe (Fig. 1e) versions of White's illusion. Blakeslee and McCourt (2004) found that WE-Howe does not lead to a consistent illusion direction across subjects, so the model's prediction of contrast actually matches what some people see. The model predicts contrast because the solid black and white horizontal bars that the test patches are on produce a strong contrast signal (note that part of the illusion is the same as SBC, an illusion the ODOG models see strongly). This contrast signal is bigger than the assimilation caused by the grating above and below the test patches, so the overall prediction is of contrast. While FLODOG predicts contrast for WE-Anderson, an illusion people consistently see as assimilation, it does predict that the illusion is closer to assimilation than is WE-Howe. This is because the test patches are offset from the contrast-inducing horizontal bars in the image. Since, however, the ODOG-based models are overly sensitive to contrast, the reduction in contrast from the offset of the test patches is still not

Table 3
Model results

	ODOG	LODOG $n = 4$	FLODOG $n = 2s$, $m = 0.5$	FLODOG $n = 4s$, $m = 0.5$	FLODOG $n = 4s$, $m = 3.0$	Illusion strength (human)
WE-thick	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	1
WE-thin-wide	**2.08**	**2.31**	**2.52**	**2.07**	**1.72**	1.1
WE-dual	−0.30	**1.11**	**1.93**	**1.67**	**1.58**	
WE-Anderson	−0.15	−0.25	−0.43	−0.03	−0.22	*1.54*
WE-Howe	−0.42	−0.47	−0.94	−0.27	−0.47	*0*
WE-zigzag	−0.51	−0.57	**1.26**	**0.91**	−0.28	
WE-radial-thick-small	−0.67	−0.55	**0.46**	**0.49**	−0.36	
WE-radial-thick	−0.41	−0.29	**0.18**	**0.18**	−0.38	
WE-radial-thin-small	−0.34	−0.20	**2.74**	**2.00**	**0.34**	
WE-radial-thin	−0.22	**0.05**	**3.24**	**2.31**	**0.43**	
WE-circular1	−0.82	−1.00	**0.28**	**0.49**	−1.36	
WE-circular0.5	−0.53	−0.65	**1.84**	**1.45**	−0.75	
WE-circular0.25	−0.38	−0.48	**3.64**	**2.58**	−0.07	
Grating induction	**2.03**	**1.77**	**0.66**	**0.41**	**1.32**	1.49
SBC-large	**4.75**	**6.33**	**3.96**	**2.37**	**6.35**	2.72
SBC-small	**6.22**	**9.19**	**5.96**	**4.01**	**10.27**	4.73
Todorovic-equal	−0.36	−0.37	**0.08**	−0.12	−0.18	0.53
Todorovic-in-large	**0.49**	**0.52**	**0.39**	**0.38**	**0.67**	0.57
Todorovic-in-small	**0.80**	**0.86**	**1.08**	**0.71**	**1.32**	1.05
Todorovic-out	**0.35**	**0.40**	**0.03**	**0.15**	**0.28**	0.37
Checkerboard-0.16	**1.10**	**0.97**	**8.03**	**6.13**	**1.36**	*1.78*
Checkerboard-0.94	**0.40**	**0.35**	−4.89	−4.05	**0.05**	*0.68*
Checkerboard-2.1	**0.69**	**0.59**	−1.48	−1.49	**0.19**	*1.36*
Corrugated Mondrian	**0.95**	**0.73**	**0.12**	−0.25	**0.09**	2.6
Benary cross	**0.09**	**0.05**	**0.05**	**0.03**	**0.06**	2.2
Todorovic Benary 1–2	−0.12	**0.54**	**0.11**	**0.10**	**0.34**	2.86
Todorovic Benary 3–4	−0.12	**0.55**	**0.14**	**0.11**	**0.36**	2.28
Bullseye-thin	−0.74	−0.56	**0.54**	**1.17**	**0.45**	
Bullseye-thick	−0.77	−0.58	**0.07**	**0.80**	**0.50**	

Illusions are listed in the same order as in Fig. 1 and Tables 1 and 2. Cells in bold indicate the model predicts that the illusion goes in the same direction people typically see it. Note that human values have been scaled so that 1.0 equals the average strength of WE-thick.

enough for the assimilation caused by the grating to dominate the overall prediction.

For SBC (Figs. 1o and p), the prediction FLODOG makes depends on the size of the normalization window, with larger windows predicting smaller illusions strengths. The original ODOG model predicts a much stronger SBC illusion than people tend to see, so the smaller prediction of the FLODOG model is more realistic, and reason to prefer the model with a larger normalization window. FLODOG accounts for the Todorovic SBC illusions (Figs. 1q–t) as well, except for Todorovic-equal (Fig. 1q). FLODOG with a 2*s* normalization window actually predicts it in the correct direction, but with a very small strength. As mentioned earlier, however, psychophysical measurements of this illusion have produced conflicting reports of which direction it goes, so it is unclear what the proper model output should be.

The FLODOG models with $m = 0.5$ do poorly on the checkerboard illusion (Figs. 1u–w). They predict that Checkerboard-0.16 (Fig. 1u) is much stronger than it really is, and predict the other two checkerboard illusions in the wrong direction. It is worth noting, however that the checkerboard illusion depends on spatial scale, and it switches from assimilation at small scales (i.e., Checkerboard-0.16) to contrast at larger scales, with the actual crossover point varying between subjects (Blakeslee & McCourt, 2004). The FLODOG model fails because it predicts assimilation at all these scales, and thus it could be failing because it has a different crossover point between assimilation and contrast. The trend (decreasing assimilation with increasing scale) is in the correct direction.

The same FLODOG models also have difficulty with the corrugated Mondrian (Fig. 1x). The 2*s* model predicts the illusion in the right direction, but predicts that it is much weaker than people see it, and the 4*s* model predicts the illusion in the opposite direction. While this does not support the model, the Mondrian stimuli may depend on more high-level factors that cannot be captured by a low-level model. Although ODOG does make a better prediction, given that the ODOG model is clearly incomplete, it is possible that its account of the Mondrian is erroneous.

The FLODOG models also predict both the Benary cross (Fig. 1y), and also the Todorovic version (Fig. 1z), which our implementation of the original ODOG model cannot account for. Note, however, that the strength of the illusion is predicted to be weaker than people see it. Finally, FLODOG predicts the Bullseye illusion (Figs. 1aa and bb), with the 4*s* version predicting a stronger illusion than the 2*s* version.

FLODOG can be made more similar to the LODOG model by setting the weighting of nearby frequencies (m) to a larger number, such as 3.0. With this parameter setting, energy at any scale within an orientation will influence the normalization of each filter. Table 3 shows how

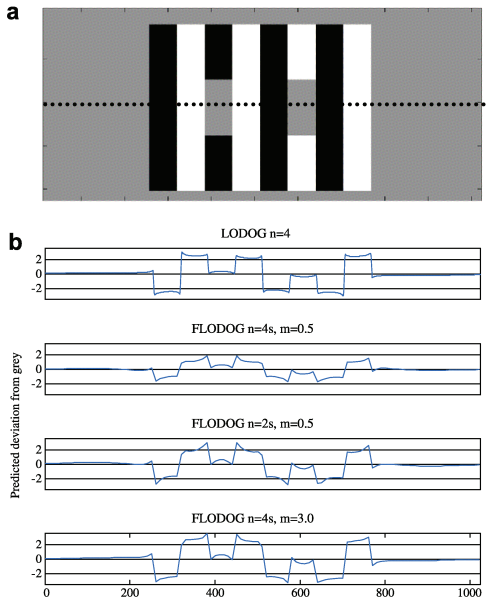


Fig. 5. Model predictions for WE-Thick. (a) Dotted line indicates location of cross-sections. (b) Perceived brightness predicted by LODOG and FLODOG models along the cross-section.

such a model performs. In contrast to the other FLODOG configurations, only the thin variants of WE-radial are predicted, no version of WE-circular is predicted, and WE-zigzag is not predicted. This configuration of FLODOG does, however, predict all versions of the checkerboard illusion, and weakly predicts the corrugated Mondrian illusion. Thus, we can see that a large part of FLODOG's success is due to normalizing each filter response primarily by itself, rather than by all the filters of the same orientation.

Cross-sections of the FLODOG model, shown in Fig. 5 for WE-thick, reveal that FLODOG predicts marked non-uniformity within regions of the input that are uniformly shaded. While no psychophysical experiments exist which directly contradict these variations, it is clear that the magnitude predicted by the model is larger than experienced when just looking at the input images. It is possible that much smaller-scale non-uniformities are measurable under psychophysical testing, and indeed Blakeslee and McCourt (1999) collected psychophysical measurements that showed the gray test patches in WE-thick do have non-uniform brightness, in the same direction as seen by the ODOG model. In addition, Blakeslee and McCourt (1997) found that the test patches in SBC and grating induction also have non-uniform brightness, of the same pattern ODOG predicts. At least for the grating induction illusion

(Fig. 1n), the non-uniform brightness predicted by ODOG is clearly visible to the untrained eye, without any psychophysical testing required. It is not known why this occurs with grating induction, and not other illusions. In any case, the non-uniform brightness perception predicted by FLODOG is not entirely wrong, though the magnitude is clearly too large.

The issue of non-uniform brightness responses cuts across different filter-based approaches, and deserves further study. While FLODOG and LODOG predict larger non-uniform responses than ODOG does for many illusions, it is worth noting that ODOG can also predict very non-uniform perceived brightness in some situations, such as SBC-large (Blakeslee & McCourt, 1999). One possibility is that these non-uniformities are only present in the early stages of brightness processing, and at some later stage the different brightness values are averaged within a region to produce a single perceived shade of gray, as suggested by Grossberg and Todorovic (1988).

In conclusion, the FLODOG model predicts many illusions that LODOG does not. The exact parameter values are not critical, but in general the quality of the predictions is better when a larger window is used for normalization (such as 4s) and when each scale is normalized relatively independently (such as when $m = 0.5$). While different parameter settings do allow the model to account for other illusions, these settings predict the widest range of illusions and are the recommended values to use when testing the model in future work.

It seems that the biologically plausible changes to LODOG implemented in FLODOG do improve the predictive power of the model. FLODOG's failure to predict illusions that are correctly predicted by ODOG and LODOG raise the possibility that the success of these models depends on less plausible mechanisms (such as normalizing each filter response equally by all frequencies of the same orientation, and a fixed size normalization window). The true cause of these effects may be due to other mechanisms that operate before or after the steps modeled by FLODOG.

6. Conclusions

This paper explored the response normalization mechanism of the original ODOG model, and found that it could be extended to be both more neurally plausible and more effective at predicting brightness illusions.

Our simplest extension, LODOG, does not reduce the functionality of the ODOG model, but only minimally increases the number of illusions correctly predicted. At the least, this shows that the response normalization process can be successfully calculated locally.

Our more advanced model, FLODOG, is not only more plausible, but also increases the number of illusions correctly predicted. The success of this model suggests that many brightness illusions could be due to low-level mechanisms in early visual processing. For instance, an area like V1 could behave much like FLODOG if lateral interactions or feedback cause cells that respond to the same orientation and similar frequencies to inhibit each other. Indeed, Rossi and Paradiso (1999) have shown that there are a significant number of cells in V1 which respond to the perceived brightness of stimuli, instead of the actual luminance (though this has only been tested for SBC-like stimuli), which supports the idea that brightness perception could occur in early visual areas. Since FLODOG uses filters of much larger spatial extent than does V1, it is clear that FLODOG is not a model of V1 at the level of individual neurons, but might represent the combined activity of groups of neurons, either in V1, or distributed across multiple early visual areas.

Why might early visual areas perform the kind of calculation that FLODOG models? Perhaps it is part of the computational processing that produces lightness constancy. Another possibility is that it is a side effect of an entirely different calculation. Schwartz and Simoncelli (2001) have developed a model of the firing rate of neurons in V1, in which neurons are inhibited by neighboring neurons that have correlated variance of firing rate. When trained on natural scenes, where adjacent regions have similar orientations and spatial frequencies, the normalization resulting from this model weights nearby spatial frequencies and orientations most heavily, similar to FLODOG. The upside of such a model is that it makes the population activity more statistically independent, when exposed to a natural image. Increasing statistical independence is thought to produce a more optimal neural code (Barlow, 1961).

There may be additional modifications to the FLODOG model that could improve its performance, without radical changes to its mechanisms. One extension that we have considered is normalizing each orientation by nearby orientations, also using a Gaussian weighting analogous to how we currently weight nearby frequencies. Experiments with this extension, however, found that normalizing across nearby orientations does not improve the predictions of the model. Another alternative, which we have not explored, is that the orientation tuning of the model might be too broad, and that more than just six orientations should be used.

FLODOG is only a model of early stages of brightness processing. There is clearly a need for some form of anchoring to explain the fact that the lightest surface in a scene tends to look white (Gilchrist et al., 1999). In addition, the percept of transparency can change whether a surface looks white or black (Anderson & Winawer, 2005). A model like FLODOG cannot explain either of these types of effects. Our work does, however, suggest that low-level mechanisms could be a significant factor in many of the illusions studied here. By itself, however, the existence of a successful low-level model does not prove that higher-level mechanisms do not contribute as well. Further work will be necessary to develop variations of these illusions that pit aspects of the high or low-level theories against each other, to determine their relative contributions.

Finally, we would like to stress the utility of having a model that can be tested on an arbitrary input image with minimal assumptions.[2] Any grayscale image can be fed into the LODOG and FLODOG models, and a prediction of brightness produced. We look forward to applying the models to new illusions to see if our low-level approach can account for other brightness illusions not studied here.

Acknowledgments

We thank Micah Richert for his role in our first implementation of the ODOG model, and Sophie Soong for implementing many of the illusions we tested. This material is based upon work supported by the National Science Foundation under NSF Career Grant No. 0133996 to VR de Sa. AE Robinson and PS Hammon were supported by NSF IGERT Grant #DGE-0333451 to GW Cottrell.

References

Adelson, E. H. (1993). Perceptual organization and the judgment of brightness. *Science, 262*, 2042–2044.

Anderson, B. L. (2001). Contrasting theories of White's illusion. *Perception, 30*, 1499–1501.

Anderson, B. L., & Winawer, J. (2005). Image segmentation and lightness perception. *Nature, 434*, 79–83.

Anstis, S. (2003). White's effect in brightness & color. Online Demonstration. http://psy.ucsd.edu/~sanstis/WhitesEffect.htm.

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). Cambridge, MA: MIT Press.

Benary, W. (1924). Beobachtungen zu einem experiment uber helligkeitskontrast. *Psychologische Forschung, 5*, 131–142.

Bindman, D., & Chubb, C. (2004). Brightness assimilation in bullseye displays. *Vision Research, 44*, 309–319.

Blakeslee, B., & McCourt, M. E. (1997). Similar mechanisms underlie simultaneous brightness contrast and grating induction. *Vision Research, 37*, 2849–2869.

Blakeslee, B., & McCourt, M. E. (1999). A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Research, 39*, 4361–4377.

Blakeslee, B., & McCourt, M. E. (2001). A multiscale spatial filtering account of the Wertheimer-Benary effect and the corrugated Mondrian. *Vision Research, 41*, 2487–2502.

Blakeslee, B., & McCourt, M. E. (2004). A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization. *Vision Research, 44*, 2483–2503.

Blakeslee, B., Pasieka, W., & McCourt, M. E. (2005). Oriented multiscale spatial filtering and contrast normalization: a parsimonious model of brightness induction in a continuum of stimuli including White, Howe and simultaneous brightness contrast. *Vision Research, 45*, 607–615.

Clifford, C. W. G., & Spehar, B. (2003). Using colour to disambiguate contrast and assimilation in White's effect. *Journal of Vision, 3*, 294a.

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

Gilchrist, A. (2006). *Seeing in Black and White*. New York: Oxford University Press.

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., et al. (1999). An anchoring theory of lightness perception. *Psychological Review, 106*, 795–834.

Grossberg, S., & Todorovic, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics, 43*, 241–277.

Hong, S. W., & Shevell, S. K. (2004). Brightness contrast and assimilation from patterned inducing backgrounds. *Vision Research, 44*, 35–43.

Howe, P. D. L. (2001). A comment on the Anderson (1997), the Todorovic (1997), and the Ross and Pessoa (2000) explanations of White's effect. *Perception, 30*, 1023–1026.

Howe, P. D. L. (2005). White's effect: removing the junctions but preserving the strength of the illusion. *Perception, 34*, 557–564.

Issa, N. P., Trepel, C., & Stryker, M. P. (2000). Spatial frequency maps in cat visual cortex. *Journal of Neuroscience, 20*, 8504–8514.

McCourt, M. E. (1982). A spatial frequency dependent grating-induction effect. *Vision Research, 22*, 119–134.

Pessoa, L., Baratoff, G., Neumann, H., & Todorovic, D. (1998). Lightness and junctions: variations on White's display. *Investigative Ophthalmology and Visual Science (Supplement), 39*, S159.

Rossi, A. F., & Paradiso, M. A. (1999). Neural correlates of perceived brightness in the retina, lateral geniculate nucleus, and striate cortex. *Journal of Neuroscience, 19*, 6145–6156.

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

Sullivan, T.J., & de Sa, V.R. (2003). Unsupervised learning of Complex Cell Behavior. Unpublished Tech Report. http://www.sullivan.to/Sullivan_CCell1.pdf.

Todorovic, D. (1997). Lightness and junctions. *Perception, 26*, 379–395.

Tootell, R. B., Silverman, M. S., & De Valois, R. L. (1981). Spatial frequency columns in primary visual cortex. *Science, 214*, 813–815.

White, M. (1979). A new effect of pattern on perceived lightness. *Perception, 8*, 413–416.

[2] Unfortunately, implementing these models is not a trivial task. To aid further work on brightness perception we will make our code (MATLAB) available for research purposes to anybody that asks. In addition, we will make available compiled versions of these models so that access to MATLAB is not required.

Chapter 1, in full, is a reprint of the material as it appears in Robinson, AE, Hammon, PS, & de Sa, VR. (2007). Explaining brightness illusions using spatial filtering and local response normalization. *Vision Research*, 47, 1631-1644. The dissertation author was the primary investigator and author of this paper.

Vision Research 48 (2008) 2370–2381

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Brief presentations reveal the temporal dynamics of brightness induction and White's illusion

Alan E. Robinson *, Virginia R. de Sa

Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, USA

ARTICLE INFO

Article history:
Received 28 February 2008
Received in revised form 26 July 2008

Keywords:
Brightness
Filling-in
Temporal dynamics
Induction
Masking

ABSTRACT

We measured the timecourse of brightness processing by briefly presenting brightness illusions and then masking them. Brightness induction (brightness contrast) was visible when presented for only 58 ms, was stronger at short presentation times, and its visibility did not depend on spatial frequency. We also found that White's illusion was visible at 82 ms. Together, these results suggest that (1) brightness perception depends on the surrounding context, even at very short presentation times, (2) the initial brightness percept is generated very quickly, but additional exposure can modulate it, and (3) the temporal dynamics are not dependent on a slow filling-in process.

Published by Elsevier Ltd.

1. Introduction

The perceived brightness of a surface depends on the brightness of the surfaces that surround it. This is known as brightness induction. One particularly well-known example of brightness induction is brightness contrast, in which the brightness of the surrounding surfaces *induces* a shift in the brightness of the center surface such that the apparent contrast is increased. In this work we investigate the timecourse of that induction. We also investigate the timecourse of White's illusion (White, 1979), which shares many visual similarities to the configuration of brightness induction we studied, but instead elicits a reduction in overall contrast.

The temporal properties of brightness induction were first investigated systematically by Magnussen and Glad (1975) (see also Glad & Magnussen, 1972). Their subject viewed a 1° spot of constant luminance embedded in a half-circle with a 3° radius. The luminance of the half-circle was modulated in time with a square wave profile, inducing a perceived brightness change in the constant luminance spot. Their subject attempted to match the perceived brightness of this induction in a separate display. Increasing the modulation rate from 0.5 to ~5 Hz made induction appear stronger, but thereafter the strength decreased, with induction disappearing above 10 Hz.

De Valois, Webster, De Valois, and Lingelbach (1986) conducted a similar study, measuring the amount of induction in a 1° square of constant luminance, embedded in a ~3° square that was sinusoidally modulating in luminance over time. With this paradigm they found that for modulations rates of 0.5 to 2.5 Hz the strength of brightness induction was relatively constant, but above 2.5 Hz it quickly fell to nearly zero. It is unclear why these studies found such different temporal cut-offs, but they do differ in several methodological details, most notably the type of temporal modulation used (square vs. sine wave).

Rossi and Paradiso (1996) found that the temporal limits of brightness induction varied as a function of spatial frequency. Subjects viewed a grating where the luminance of every other stripe was varied sinusoidally in time, and were asked to adjust the rate of modulation to the minimum temporal frequency where no brightness induction was visible in the unmodulated stripes. For the widest stripes (16°) the thresholds were between 0.8 and 1.8 Hz, but for the thinnest stripes tested (0.5°) the threshold increased to between 1.5 and 5 Hz, depending on the subject.

They argued that this dependence on scale is consistent with the theory that brightness perception depends on the retinotopic filling-in of neural signals. According to this theory, the visual system first detects the contrast at borders between uniform regions, and then propagates this contrast information from the borders into the uniform regions. Importantly, the propagation, or filling-in, takes time which is dependent on the distance that the signal must travel.

Rossi and Paradiso (1996) ran two other experiments that were also consistent with this theory. Using fixed modulation frequencies (0.5, 1, 2, and 4 Hz), they asked subjects to make brightness matches to gratings at several different spatial frequencies. All subjects showed a reduction in illusion strength as the temporal frequencies increased from 0.5 to 2 Hz, and no illusion at 4 Hz at all. This can be explained by filling-in if one assumes that the

* Corresponding author. Fax: +1 858 534 1128.
E-mail address: robinson@cogsci.ucsd.edu (A.E. Robinson).

0042-6989/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.visres.2008.07.023

partially filled-in signals are averaged over time. In their third experiment they measured the temporal phase of the induced modulation relative to the actual modulation, and found that there was a lag in the induced modulation. The amount of phase lag increased with wider stripes. This is also consistent with filling-in, if you assume that no induction is seen until the signal is propagated from the borders all the way into the center of the stripes. Indeed, based on this theory, Rossi and Paradiso calculated that filling-in travels about 140–180°/s.

If filling-in occurs at 140–180°/s, however, that cannot be the only temporal limit on brightness induction. Consider a 0.5° stripe. At 140°/s, it would be filled-in after only 1.78 ms, which would suggest that induction should be seen at modulation rates of 280 Hz in their paradigm. Of course, due to the critical flicker fusion threshold one would expect modulation to disappear much sooner, perhaps around 50 Hz. Nonetheless this is much higher than the 1.5–5 Hz range that they found in their first experiment for stimuli of this size. Presumably, even if filling-in is involved in the temporal limits on induction, there are other factors as well.

A very different paradigm has also found evidence in support of filling-in. Paradiso and Hahn (1996) showed that steadily decreasing or increasing the luminance of a disk led to a slightly delayed change in the perceived brightness at the center of the disk. This suggests that filling-in occurs and that it does not occur instantaneously, though it still could be quite fast.

Even if filling-in does occur, there is some question as to whether it plays any role in the temporal limits of brightness perception. Davey, Maddess, and Srinivasan (1998) have reanalyzed the data from Rossi and Paradiso's first experiment where subjects adjusted the modulation rate until no induction was seen. They found that in order to explain these results a much slower speed of filling-in must be posited, between 9 and 14°/s, which is in very poor agreement with the speed Rossi and Paradiso estimated from their third experiment.

Gunther and Dobkins (2005) had subjects adjust a 3.5° disk that alternated between red and green so that the two colors appeared equiluminant, while at the same time an annulus surrounding the disk modulated between white and black. They found that the induction caused by the annulus was reduced and then disappeared when the modulation rate of the entire figure was between 8 and 20 Hz, but if the modulation rate was increased even further some subjects saw a reappearance of induction. The reappearance of induction at faster modulation rates is difficult to explain with filling-in.

There is also some evidence suggesting the temporal dynamics of the Craik-O'Brien-Cornsweet effect (COC) cannot be explained by filling-in (a general review of the COC effect can be found in Kingdom & Moulden, 1988). Devinck, Hansen, and Gegenfurtner (2007) asked subjects to select the modulation frequency at which no COC illusion was seen. For stripe widths of 10–2.5°, they found that the achromatic COC could be seen at faster modulation rates for the narrower stripe widths (a similar result was found by Davey et al. (1998)). While at first pass this is compatible with filling-in, they note that to explain their results filling-in would have to travel at slower speeds for thinner stripes. In addition, for chromatic COCs of the same widths, and for achromatic COCs with smaller stripe widths of 2.5–0.4°, they found the reverse effect. Decreasing the stripe width caused the COC to only be seen at slower modulation rates. While these results do not favor the filling-in theory, it is also possible that slow filling-in plays a role in brightness induction, but not in the COC illusion.

All of the evidence for the speed of filling-in for brightness induction and the temporal properties of brightness perception has been based on the response to temporally modulating figures. These experiments are not in good agreement with each other. Some of the evidence against filling-in comes from the chromatic domain, which raises the possibility that filling-in occurs only in the luminance domain, though we feel this is unlikely. One would expect some difference between chromatic and achromatic stimuli, since the two are processed by different mechanisms, but it would be surprising if one requires filling-in and the other does not.

Another difficulty with the previous experiments is that they were based on the response to modulation over the period of at least several seconds. Thus, they cannot be used to determine the timecourse of perception relative to the initial onset of the figure. For these reasons we elected to explore the temporal dynamics of brightness induction using a new paradigm.

In our work we investigated the timecourse of induction by having subjects make brightness matches to a briefly presented static stimulus. To limit processing time after the stimulus was removed we covered it with a noise mask. We reasoned this paradigm would have several advantages. First, all components of the stimulus have the same duration of exposure to the subject, rather than being a combination of modulating and constant regions. Second, by showing a stimulus and then masking it, there is less potential for competing precepts. In the Rossi and Paradiso paradigm the stimulus alternates between two opposite percepts. How these competing percepts are resolved might influence what is perceived while having little to do with the timecourse of brightness perception itself.

Our paradigm is somewhat similar to the work of Paradiso and Nakayama (1991), which studied brightness percepts elicited by briefly presenting a large white disk and then masking it with a smaller pattern, such as a black circle with a white outline. The key difference is that in the Paradiso & Nakayama work, the mask consisted of a pattern and not a noise mask. Thus, rather than trying to stop additional processing, they were studying how the pattern might influence the perception of the previous figure.

2. Experiment 1

Based on Rossi and Paradiso's results we expected two effects: First, as the presentation of the stimulus was made shorter and shorter the illusion strength should decrease. If the presentation time was short enough, the illusion should disappear. Second, the point at which the illusion should disappear would occur at different times depending on the spatial frequency of the induction stimulus; a lower spatial frequency stimulus should give rise to induction only when displayed for a relatively longer period of time than a higher spatial frequency stimulus. To test these two predictions we used two different spatial frequencies and a range of presentation times.

2.1. Methods

2.1.1. Subjects

One author and three naive subjects participated in the experiment. The naive subjects had varying levels of psychophysics experience, but none had prior experience with brightness matching experiments, or the hypothesis being tested.

2.1.2. Apparatus

Stimuli were presented on a 21″ NEC FE2111SB CRT driven by an ATI RADEON 7000 VE video card at a refresh rate of 85 Hz. Display luminance was linearized using a color lookup table that drove a 10-bit DAC over a range of 0 to 102 cd/m^2. A Cambridge Research Systems ColorCal colorimeter was used to select the appropriate lookup table values. A chinrest was used to maintain a viewing distance of 72 cm. Stimuli were generated and displayed using Matlab running the Psychophysics Toolbox, version 2.54 (Brainard, 1997; Pelli, 1997). The experiment was run in a dark room and subjects adapted to the light level for 3 min before collecting data. The same apparatus was used in all experiments.

2.1.3. Stimuli and procedure

We measured the strength of induction as a function of how long the stimulus was displayed before being replaced with a mask (*OnTime*, 58, 82, 117, or 1120 ms), using the method of adjustment. A diagram of the procedure is shown in Fig. 1. The stimulus was a grating made up of *inducing* stripes and *target* stripes. Each *inducing* stripe had the same luminance on a given trial, either 12 cd/m^2 (*gray*) or 102 cd/m^2 (*white*). Each *target* stripe had the same luminance on a given trial, either 51, 57, or 64 cd/m^2. We designed the stimulus so that when it was displayed every region was an increment relative to the pre-stimulus blank screen. This ensured that all regions of the figure triggered transient responses, rather than a combination of steady state and transient responses. In the high-frequency condition each stripe in the grating was 1° wide and 12° tall; in the low frequency condition each stripe was 10.6° by 12°.

Trials were grouped into conditions; within a condition we held constant the *inducing* stripe luminance, the stimulus *OnTime*, and the spatial frequency of the induction grating. We varied the *target* stripe luminance (51, 57, or 64 cd/m^2) between trials to make sure subjects were attending to the *target* stripe and not answering based on memory from previous trials or conditions. Subjects completed 12 trials per condition and on average repeated each condition 6 times.

For each trial the following three frames were shown in a loop until the subject had completed making the match: (1) a pre-stimulus blank (0 cd/m^2) for 306 ms, (2) the induction stimulus, and (3) a noise mask that exactly covered the induction stimulus for 894 ms. The noise mask was constructed out of 0.25° squares. On even lines the luminance of each square was selected at random from a uniform distribution from 0 to 102 cd/m^2. On odd lines, each square was set to 102 cd/m^2 minus the luminance of the square above it. Thus, the space-averaged luminance of each vertical pair of squares was 51 cd/m^2.

The induction stimulus was displayed on the lower half of the screen. To quantify the strength of induction we had subjects adjust the luminance of a constantly visible patch on the upper half of the screen to match the appearance of the briefly presented induction stimulus. The patch was set to a random luminance at the beginning of each trial (between 0 and 102 cd/m^2, uniformly distributed), and was adjustable in 0.4 cd/m^2 increments using the computer's keyboard. The adjustment patch was 1° by 2° with a black and white checkerboard border with each check covering 0.5° × 0.5°. The subject's task was to match the subjective shade of gray of the *target* stripe at the center of the screen (a small dot was placed above this stripe to orient subjects). The induction stimulus was displayed multiple times in a trial so that subjects could continue to make adjustments until they were satisfied.

2.1.4. Results

We averaged over the matches subjects made for different *target* stripe luminances to obtain an overall measure of induction at different *OnTimes*, for the two different stripe widths. In Fig. 2 we show separate curves for the conditions where the *inducing* stripes were bordered by *gray* and *white*, respectively. The vertical distance between the curves is the overall strength of induction.

For the 1° wide stripes we found that decreasing *OnTime* increased the strength of induction in 3 out of 4 subjects. Over the range of *OnTimes* tested, there was no evidence that induction went away at any *OnTime*. At the shortest *OnTime* tested (58 ms), however, subjects complained that they could no longer reliably resolve the difference between *target* and *inducing* stripes. Thus we did not collect any brightness matches for this presentation time. These results suggest that so long as the stimulus is discernable, brightness induction occurs. This is in clear contrast to Rossi and Paradiso's conclusion that induction gets weaker at higher speeds, and that it eventually disappears.

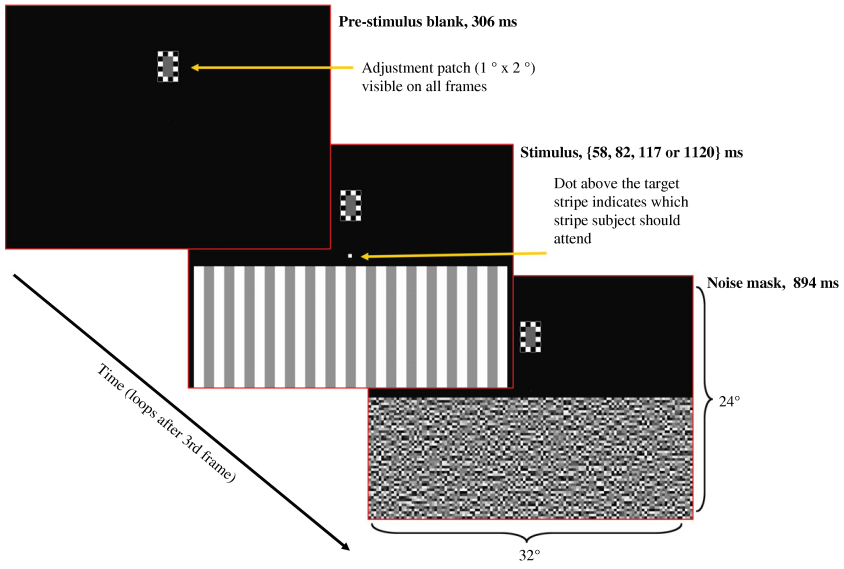


Fig. 1. Diagram of experimental paradigm. The three frames illustrated here are shown in a loop until the subject is satisfied that the adjustment patch matches the target stripe. The duration of the stimulus is constant within any block of trials.

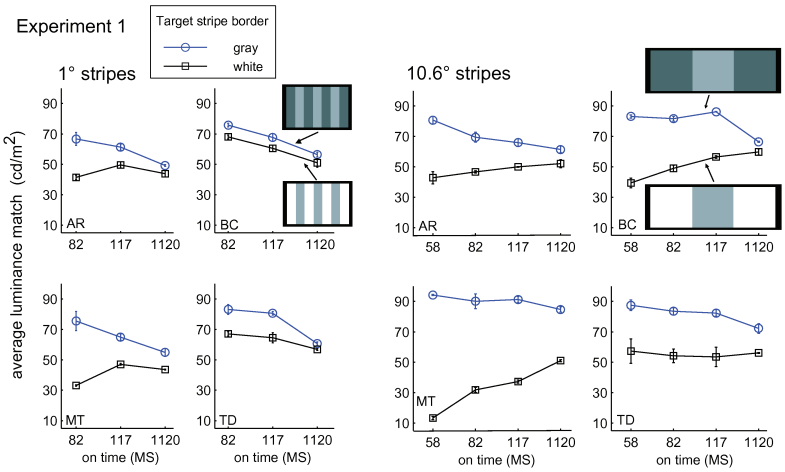


Fig. 2. Experiment 1. Average luminance matches for the target stripe. The vertical distance between the border gray and border white curves is the strength of the brightness induction illusion. The insets show schematic illustrations of the stimuli presented (*Note:* The illustrations are not to scale, and the number of bars shown is fewer than in the actual 1° stimuli). Error bars denote standard error of the mean.

For the 10.6° wide stripes, we found similar overall results. Decreasing *OnTime* consistently led to an increase in the strength of induction in all subjects. Even at the shortest *OnTime* we tested (58 ms) subjects were able to make brightness matches. This is in clear contrast to the slow filling-in theory, which would suggest that induction should be slower for the wider stripes. Our results suggest little difference as a function of stripe width, and if anything it is possible to make brightness matches at even higher speeds with wider stripes.

One potential concern is that subjects might not have been able to see the stimuli clearly in all conditions (especially at the shortest *OnTimes*) potentially introducing some memory-bias in their brightness matches. To address this we also analyzed how well subjects' brightness matches tracked the actual luminance variation in the *target* stripes that occurred between trials. This data is shown in Fig. 3 for *target* stripes of luminance 51 and 64 cd/m^2 (57 cd/m^2 is omitted to reduce clutter). While the individual matches were variable, we found that the mean of the subjects' matches did vary as a function of the luminance of the *target* stripe for *OnTimes* of 1120 ms and 117 ms, for both stripe widths. For shorter *OnTimes* some subjects began to have difficulty, suggesting that accurate luminance perception began to fall apart, even though induction still appeared to occur. This could be because subjects were making matches based on memory, but it could also be because the noisy nature of perception at these high speeds makes it difficult to distinguish small luminance differences. We tested this in a second experiment.

3. Experiment 2

In Experiment 1 we could not be sure that subjects were making brightness matches based on the appearance of the *target* stripe on each trial. To make this easier to detect we used a larger difference in actual luminance of the *target* stripe between trials. Furthermore, in the previous experiment we did not collect data at the shortest *OnTime* for the 1° stripes, because we did not want to encourage subjects to guess when they felt that they could not see the stimulus clearly. In Experiment 2 we instead asked subjects to make their best guess, even if they felt unsure.

3.1. Methods

3.1.1. Subjects

The same subjects from Experiment 1 participated in Experiment 2. Each subject completed Experiment 1 before starting Experiment 2.

3.1.2. Stimuli and procedure

The stimuli and procedure was the same as Experiment 1, with the following modifications: We only used two *target* stripe luminances (31 or 72 cd/m^2), and we only collected data for 82 and 58 ms *OnTimes*.

3.1.3. Results

In all conditions subjects' responses varied as a function of the *target* stripe luminance, shown in Fig. 4 as the vertical distance between the dashed and solid lines. This demonstrates that subjects can perceive the target stripes at even the shortest *OnTime*, and are basing their responses on the appearance of those stripes on each trial, and not on their memory of previous trials where *OnTime* was longer. Furthermore, in all conditions, subjects saw strong induction effects including when the stripe width was 1° and *OnTime* was only 58 ms. This is shown Fig. 4 as the vertical offset between pairs of dashed lines, or sold lines. This provides further evidence that the amount of presentation time necessary to see induction does not depend on the spatial frequency of the stimulus.

4. Experiment 3

In the previous two experiments we used a brightness mask of a higher spatial frequency than the brightness induction grating. From the perspective of filling-in, this higher-frequency mask

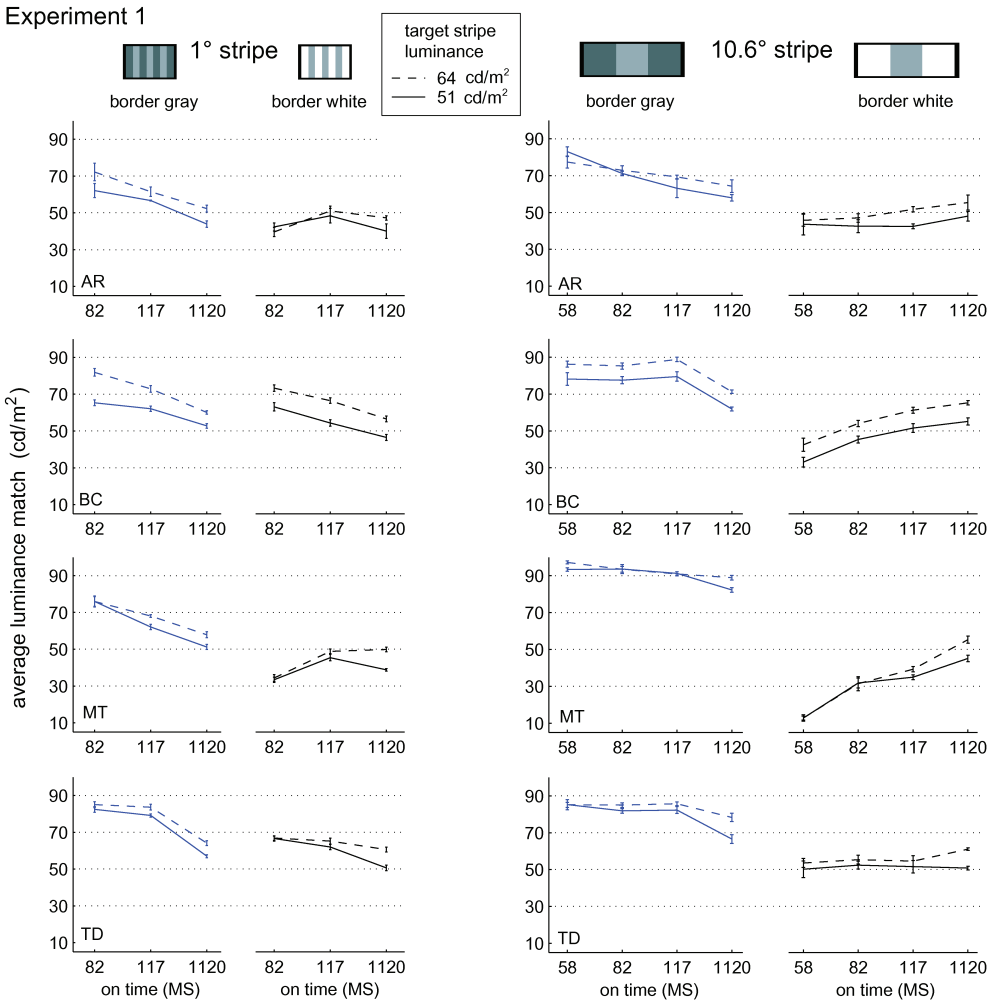


Fig. 3. Experiment 1. Average luminance matches for the target stripe on trials where the target stripe was 64 cd/m2 (dashed lines) or 51 cd/m2 (solid lines). The distance between the curves in each column shows how well subjects tracked the relative difference in target stripe luminance. The distance between curves in pairs of columns shows the strength of brightness induction. The insets show schematic illustrations of the stimuli presented (*Note:* The illustrations are not to scale, and the number of bars shown is fewer than in the actual 1° stimuli). Error bars denote standard error of the mean.

should be ideal for stopping additional processing, since it contains many edges. But if the temporal dynamics of brightness perception are not limited by filling-in, it may not be the ideal mask. According to spatial filtering theories of brightness perception, brightness induction is due to contrast-sensitive spatial filters tuned to the spatial frequency of the inducing grating. Thus, it is possible that the higher-frequency mask we used stopped processing of high-frequency information, but did not disrupt processing of the lower frequencies that actually caused induction. In Experiment 3 we used a potentially more effective mask, with the same spatial frequency as the horizontal spatial frequency of the induction grating.

We first implemented this in the paradigm used in Experiments 1 and 2, using an *OnTime* of 58 ms. Interestingly, we found that the first time the induction stimulus was presented in a trial it was not too difficult to see, but after viewing the mask, additional presentations of the induction stimulus were invisible. We theorize that this is due to frequency-specific visual adaptation to the mask,

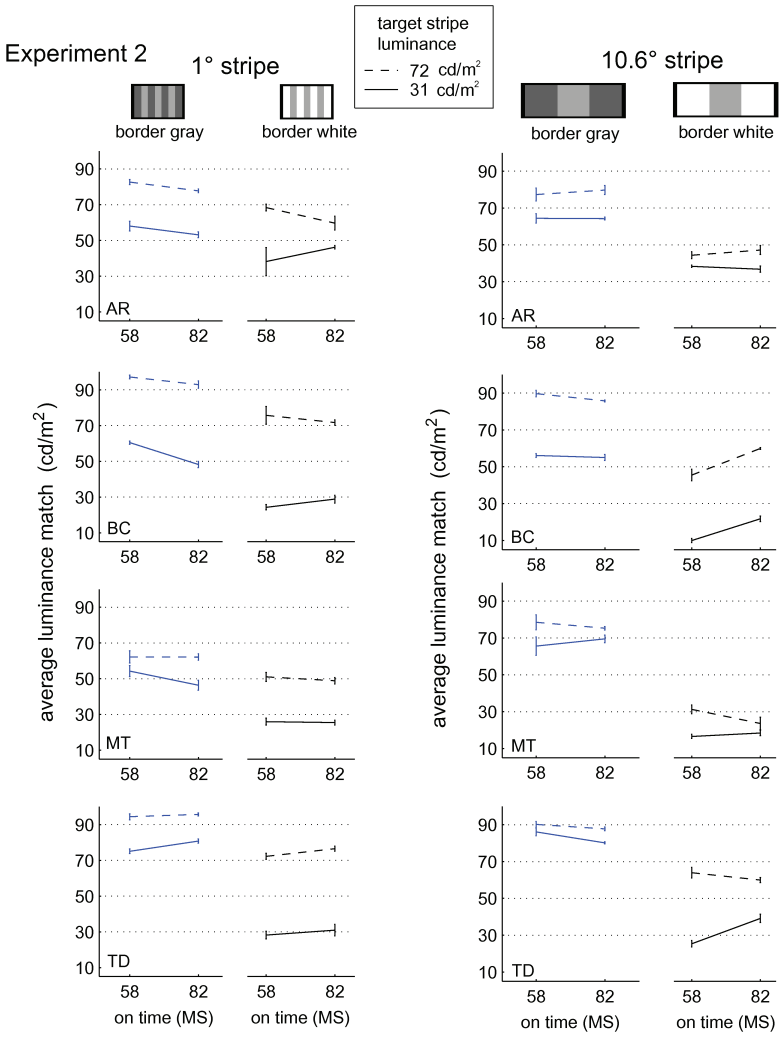


Fig. 4. Experiment 2. Average luminance matches for the target stripe on trials where the target stripe was 72 cd/m2 (dashed lines) or 31 cd/m2 (solid lines). The distance between the curves in each column shows how well subjects tracked the relative difference in target stripe luminance. The distance between curves in pairs of columns shows the strength of brightness induction. The insets show schematic illustrations of the stimuli presented (*Note:* The illustrations are not to scale, and the number of bars shown is fewer than in the actual 1° stimuli). Error bars denote standard error of the mean.

which led to an overall reduction in sensitivity to any stimulus made up of those spatial frequencies. To prevent adaptation, we elected to use a new paradigm where the induction stimulus was presented *just once* in each trial, and then covered by a mask, which was shown until the subject had made a brightness match. After inserting a short break between trials we found that now subjects could make as many brightness matches as we asked in a single session. This paradigm, while noisier because subjects do not have the chance to check their match against the induction stimulus, has the additional advantage that there is no opportunity to integrate information from multiple presentations of the stimulus. We feel it is unlikely that this occurred in the first 2 experiments, since the appearance of the induction grating seemed constant across the multiple presentations within a trial. Nonetheless, using one presentation per trial protects against this potential problem.

4.1. Methods

4.1.1. Subjects

Four subjects participated. One (AR) was an author, and one (JB) was familiar with the general purpose of the experiment. The other two subjects had minimal psychophysics experience, other than participating in Experiments 1 and 2, and were naive to the purpose of the experiments.

4.1.2. Stimuli and procedure

We replicated the 1°-wide strip induction stimulus from Experiment 2 with some minor changes. We found that with an *OnTime* of 58 ms and only one exposure to the stimulus it was difficult to detect which stripe was the target stripe, even though it was clear that the grating was made up of dark and light stripes. Therefore, we doubled the number of trials and asked subjects to make brightness matches to both the dark stripes and the light stripes. At the beginning of each trial subjects were told which stripe to match. After 2 s this prompt disappeared and subjects fixated a dot, centered in the lower half of the screen, where the grating was to appear. After an additional 1.5 s the induction grating was displayed for 58 ms (*OnTime*). The odd stripes of the grating were either 12 or 102 cd/m^2 (corresponding to the *border gray* and *border white* conditions of Experiments 1 and 2), and the even stripes were 19, 31, or 72 cd/m^2. Note that the odd and even stripes no longer consistently map onto the *inducing* and *target* stripe terms used in Experiments 1 and 2, since either the odd or even stripes could be targets, depending on the trial. All brightness combinations were tested in random order. Since brightness matches were made to both dark and light stripes, each combination was repeated twice within a block of trials. Subjects completed 16 blocks of trials on average.

The mask was shown next. It differed from Experiments 1 and 2 only in that each square was 1° wide. The adjustment patch was displayed in the upper half of the screen, with the same checkerboard border as used in Experiments 1 and 2. Instead of using the keyboard to adjust the luminance of the patch, however, subjects used a mouse, with movements to the left darkening the patch and movements to the right brightening it. The initial luminance of the patch was set randomly on each trial, so the relationship between absolute mouse location and brightness also changed on each trial. Using the mouse made it much easier to select a matching brightness quickly before the memory of the inducing stimulus faded, though it made very fine brightness adjustments somewhat more difficult.

When the subject was satisfied with the match they would click the mouse to continue to the next trial. We also allowed the subject to abort the current trial if they felt they could not make a good match, such as because of blinking during the presentation of the induction stimulus. If they aborted, a different trial would be shown next, and the aborted trial would be displayed again later in the experiment. We encouraged subjects to use this feature whenever they felt they could not make a good match, for whatever reason.

In between trials subjects were given a 5 s rest to reduce any adaptation effect.

4.1.3. Results

We first consider the results that are analogous to the *border gray* and *border white* conditions in Experiments 1 and 2. This data is shown in Fig. 5a. As in the first two experiments, we found strong induction effects in the appearance of the *target* stripe (the strength of the illusion is the vertical offset between the two curves). All four subjects showed that *target* stripes bordered by *gray* (12 cd/m^2) appeared brighter than *target* stripes bordered by *white* (102 cd/m^2). Furthermore, subjects' brightness matches showed sensitivity to the actual brightness of the *target* stripe (in the figure this corresponds to the increase in match brightness from left to right), except for subject MD, who's *border gray* data was only poorly correlated with *target* stripe brightness. This shows that subjects were not just guessing at the proper brightness match based on whether they were told to match the bright or dark stripes.

The data shown in Fig. 5a replicates the effects found in Experiment 2. In addition, we collected data to determine if we could measure induction effects not just by changing the border of the *target* stripe from an increment to a decrement, but also from changing the luminance of the *border* stripe while holding the *target* stripe luminance and the polarity relationship constant. This data is shown in Fig. 5b. In this analysis, we expect the apparent brightness of the *target* stripe to decrease as we increase the brightness of the bordering stripes. In the figure, this would appear as a decrease in each curve from left to right. All subjects show clear evidence of this when the *target* stripe is a decrement; for the increment *target* stripe the effect is weaker, and subject MD shows the reverse trend. Nonetheless, when considering the increment and decrement data together, it is clear that even at 58 ms, the relative brightness of the *border* stripe matters, not just its polarity relationship to the *target* stripe.

From these analyses we conclude that brightness induction occurs with a single, 58 ms-long presentation, even when the brightness mask is ideal for preventing additional processing.

5. Experiment 4

In our final experiment we used our masking paradigm to measure the timecourse of White's illusion (White, 1979). White's illusion is a brightness illusion where one of the stripes of a black and white grating is partially replaced by a gray patch. The brightness of the gray patch appears to shift toward the brightness of the bordering stripes, which is the reverse of what happens in brightness induction. For this reason, White's illusion has been suggested by some authors to require more complex mechanisms than brightness induction (Anderson 1997; Todorovic, 1997). Thus, it might have a different timecourse then brightness induction.

5.1. Methods

5.1.1. Subjects

Four subjects participated. One (AR) was an author, and one (JB) was familiar with the general purpose of the experiment. The other two subjects had minimal psychophysics experience, and were naive to the purpose of the experiment.

5.1.2. Stimuli and procedure

The stimuli consisted of an 11° by 12° grating with 1° wide stripes that alternated between gray (12 cd/m^2) and white (102 cd/m^2). The *target patch* was placed on top of the central stripe, and was 1° by 2°. On each trial the *target patch* was randomly set to either 51, 57, or 64 cd/m^2. In the *on white* condition the grating was aligned so that the central stripe was a white stripe; in the *on gray* condition it was a gray stripe. The procedure was the same as Experiment 1, except that we used *OnTimes* of 82, 117, and 1120 ms. We did not include the 58 ms condition used earlier experiments because subjects found it very difficult to see the test patch at that short a display time.

5.1.3. Results

The brightness matches are shown in Fig. 6. The strength of the illusion is the distance between the curves for the *on white* and *on gray* conditions. Three of the four subjects saw a clear White's illu-

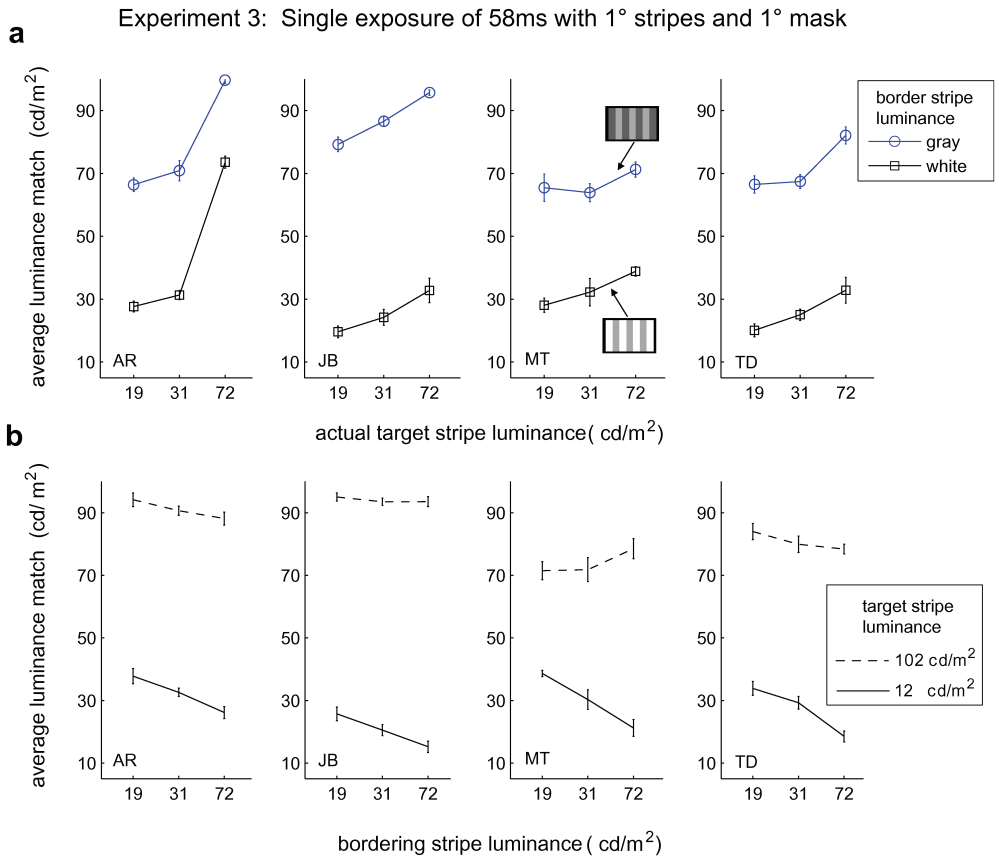


Fig. 5. Experiment 3. (a) Average luminance matches when the target stripe was bordered by gray or white. The vertical distance between curves shows the strength of the illusion. The rise in the curves from left to right shows how well subjects' tracked the actual changes in the luminance of the target stripe. The insets show schematic illustrations of the stimuli presented (*Note:* The illustrations are not to scale and fewer bars are shown than in the actual stimuli). (b) Average luminance matches as a function of increasing bordering stripe luminance. Error bars denote standard error of the mean.

sion at the shortest *OnTime* tested (82 ms). Subject SZ's data follows the same trend, but is too noisy to be conclusive. Our results suggest that White's illusion requires a similar amount of *OnTime* to become visible as does brightness induction. As in the other experiments, subject's matches also reflected the actual trial-to-trial differences in *target* patch brightness, showing that they were not using memory to make their response at the shortest *OnTime*. The average brightness matches made when the *target* patch was 51or 64 cd/m^2 is shown in Fig. 7.

The effect of *OnTime* on illusion strength was variable. In three out four subjects the strength of the illusion appeared to get slightly weaker with shorter *OnTimes*, but one subject showed the opposite trend. Due to the variability between subjects, it is unclear how the strength of White's illusion changes with longer exposure. This does suggest that something somewhat different is occurring than with brightness induction, however, where we saw a clear increase in illusion strength for the shorter *OnTimes*.

The short timecourse and minimal difference in timing between White's illusion and brightness induction is compatible with models of brightness perception that depend on the interactions of simple visual features that could be quickly computed in early visual areas. In particular, this includes models based on spatial filtering and response normalization: the ODOG (Blakeslee & McCourt, 1999) and FLODOG models (Robinson, Hammon, & de Sa, 2007), and the model of Dakin and Bex (2003). These models are compatible with our results in the sense that they predict fast brightness perception for both types of illusions, and no temporal dependence on spatial scale. The ODOG and FLODOG models do not include any explicit temporal aspect, however, so they are agnostic to the change in illusion strength we found with shorter presentations in Experiments 1 and 2. The Dakin and Bex model does apply response normalization in an iterative fashion, but the authors' claim that is only because it simplifies the implementation of the model, so it too appears to be agnostic to any temporal variation in bright-

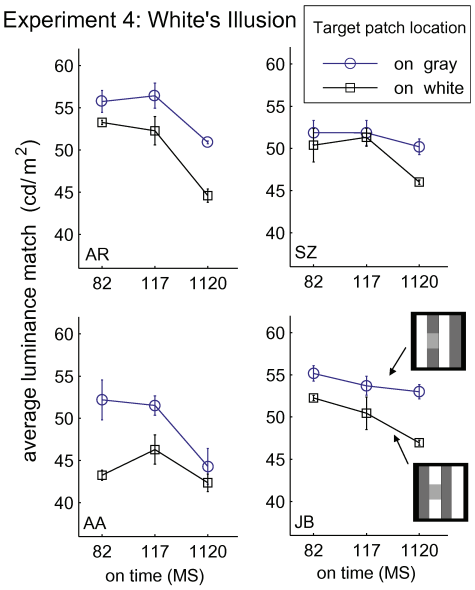


Fig. 6. Experiment 4. Average luminance matches for the target patch. The vertical distance between the on white and on gray curves is the strength of White's illusion. The insets show schematic illustrations of the stimuli presented (*Note:* The illustrations are not to scale and fewer bars are shown than in the actual stimuli). Error bars denote standard error of the mean.

ness. One possible way to modify both of these models to include a temporal aspect is to assume that response normalization completes significantly later than spatial filtering, and that the onset of the mask in our paradigm interferes with normalization. At least for the ODOG/FLODOG models, however, this would predict that brightness induction would be unchanged by varying *OnTimes*, and that White's illusion should get weaker with shorter *OnTimes*. This does not appear to match the results of our experiments, though it is true that three out of four subjects in Experiment 4 did see a reduced White's illusion with shorter *OnTimes*.

6. General discussion

The interpretation of our results depends on whether or not our masking paradigm successfully stopped additional processing of the brightness induction and White's illusion stimuli. There is some debate in the visual masking literature as to the exact effect of masking. Enns and Di Lollo (2000) suggest that rather than stopping visual processing, masking reduces target visibility at a later stage because the onset of the mask captures visual attention. This maps poorly on to our experiments, however, since according to this theory the major change with shorter *OnTimes* should be the percentage of presentations where any stimuli other than the mask was visible. Instead, we found that the brightness illusion stimulus was nearly always visible, and that changing *OnTimes* changed its appearance. In contrast, Reeves (2007) argues that visual masking does indeed stop processing, so long as the mask stimulates the same visual channels as the stimulus it is intended to mask. The definition of channel here is vague, but it can be reasonably argued

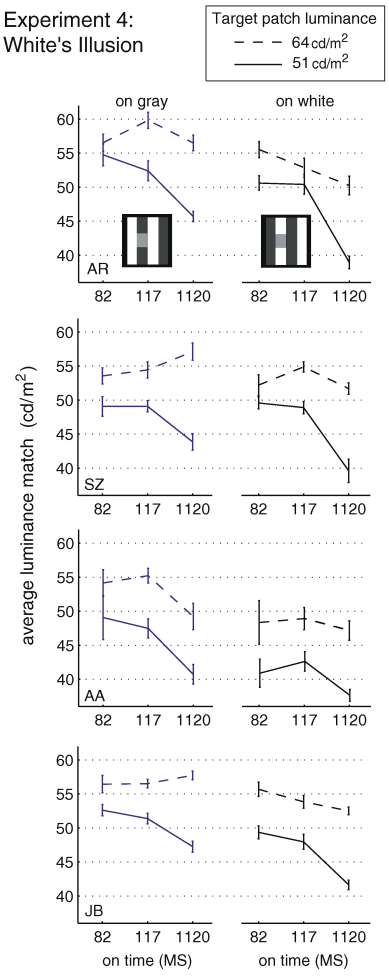


Fig. 7. Experiment 4. Average luminance matches for the target patch on trials where the target patch was 64 cd/m2 (dashed lines) or 51 cd/m2 (solid lines). The distance between the curves in each column shows how well subjects tracked the relative difference in target patch luminance. The distance between curves in pairs of columns shows the strength of White's illusion. The insets show schematic illustrations of the stimuli presented (*Note:* The illustrations are not to scale and fewer bars are shown than in the actual stimuli). Error bars denote standard error of the mean.

that our noise mask was sufficiently similar to our brightness illusion stimuli to stop additional processing. Under this assumption, our results have several implications for the speed of brightness processing and its relationship to filling-in. We will discuss each in turn.

6.1. The speed of brightness processing

First, our results suggest that brightness processing can be very fast. We showed that only 58 ms of exposure is sufficient to per-

ceive brightness induction, and 82 ms is sufficient for White's illusion. We should note that while 58 ms of exposure was sufficient to perceive brightness induction, this does not necessarily mean that a brightness percept is generated 58 ms from the onset of a stimulus. In fact translating our data into a direct measure of the speed of brightness perception requires several assumptions. If a stimulus is followed by a brightness mask after 58 ms, we assume that processing in each area is interrupted with the arrival of the mask signal. If we then assume that the stimulus and the mask that follows it are both transmitted throughout the visual system at the same rate, it follows that the processing in each area needs at most 58 ms after the signal from the retina arrives in order for a brightness percept to form (see Fig. 8). The time before the brightness percept is perceived (or speed of brightness processing) would therefore depend on the particular area that controls the perception.

There is evidence that early visual areas play a role in brightness processing. Rossi and Paradiso (1999) found that 10–30% of the cells they recorded from in cat V1 responded according to brightness percepts. In particular, they found many cells that were modulated by the brightness of flankers outside of the cells' receptive fields, in the same direction as brightness induction. Interestingly, they found that if the flankers were modulated sinusoidally in time at different frequencies, that higher speed modulation resulted in less change in the cell's firing rate, much like their previous psychophysical findings (Rossi & Paradiso, 1996). Schroeder, Mehta, and Givre (1998) report that V1 in Macaque first responds about 20–30 ms after stimulus onset. If we add this latency to our estimate of the processing time required to first develop a brightness percept, it suggests that brightness induction is visible by about 80 ms, and White's illusion by 100 ms. Interestingly, this is in rough agreement with an ERP study of White's illusion (McCourt & Foxe, 2004), which found that White's illusion induced changes in the C1 ERP component around 50–80 ms after stimulus onset.

It is important to note that we found evidence of brightness induction occurring at the very fastest speed we tested. It is very possible that induction occurs at even shorter presentation times, however in pilot work we found it very difficult to see the stimuli when we tested shorter *OnTimes*, and the brightness matches subjects made were highly variable. But it is certainly possible that brightness induction occurs at even shorter presentation times.

Our data suggests that brightness induction occurs much quicker than suggested by the data from Rossi and Paradiso (1996) second experiment where subjects made brightness matches to a temporally modulated induction stimulus. They found that induction could be measured for 2 Hz modulations, but disappeared entirely for 4 Hz modulations. At 2 Hz a full cycle from black to white and back takes 500 ms. In order for induction to be seen, however, the visual system must respond to each half cycle, (e.g. from mean gray to black and back to gray), which is 250 ms for 2 Hz, and 125 ms for 2 Hz. This would suggest that brightness induction takes less than 250 ms, but more than 125 ms. Note, however, that the sinusoidal modulation muddies the issue a little, since for part of the cycle the surround is not significantly different from gray. Induction is likely to be hard to see until the surround has reached a sufficient level of contrast relative to the center. Our technique does not suffer this ambiguity.

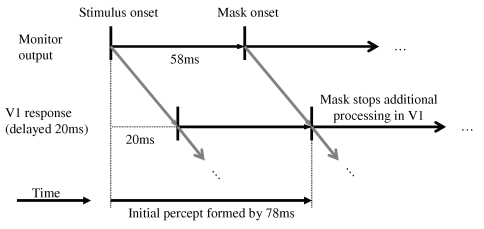


Fig. 8. Diagram of the latency of brightness processing, with respect to stimulus onset. See text for full details.

Blakeslee and McCourt (2008) recently investigated the time-course of brightness induction using the grating induction illusion, and also found evidence that induction can be seen at high modulation rates. Subjects viewed a sinusoidal grating that was modulated in counterphase. This induced a modulating grating 180° out of phase in a medium-gray test stripe that bisected the modulating grating. By adding an additional grating to the test stripe they were able to elicit movement percepts, even when the modulation of the inducing grating was as high as 24 Hz. By measuring the perceived direction of motion as a function of phase difference they were able to show that there is little to no change in phase between the induced grating and the inducing grating as the modulation rate increases. This suggests that induction sufficient to drive the motion percept occurs at the same speed as the perception of the inducing grating. Since they did not collect brightness matches, however, the exact nature of the induction percept is unclear. Together with our results, this strongly suggests that the Rossi and Paradiso's estimate of the speed of brightness perception is too slow. Blakeslee & McCourt's data, however, do not reveal the dynamics of brightness perception as a function of exposure time.

While our data shows that 58 ms of exposure is sufficient to perceive brightness induction, we also found that the strength of the illusion changed with even longer exposure. This suggests that the visual system computes the brightness of the stimulus quickly, and then refines that estimate over time. Since the illusion strength continues to change between 117 ms and 1120 ms, it appears that this refinement continues until some point after 117 ms after stimulus onset. Further work will be necessary to estimate when the brightness percept reaches a steady state.

What is the nature of this on-going processing? One possibility based on previous results is that it is a brightness filling-in signal, but we can think of no reasonable explanation as to why giving filling-in more time would lead to a weakening of brightness induction. Filling-in may occur, but it does not explain our effect. A more promising hypothesis is based on the well-known finding that the visual system tends to respond most strongly to the onset of a stimulus. For instance, consider the response of cells in Monkey V1. Albrecht, Geisler, Frazor, and Crane (2002) measured the temporal response of V1 cells to gratings of different contrasts. They found that neurons tended to reach peak firing rates around 50 ms after stimulus onset, and then firing would decay significantly to a sustained level, typically around 100 ms. Note, however, that some cells took much longer to reach a sustained level, and the firing rate histograms were quite varied on the whole. Higher contrast gratings would also increase the firing rate. Thus, the early firing rate for a low-contrast grating is similar to the later, sustained response to a high-contrast grating. If the visual system does nothing to correct for this bias, and only used the rate of firing to judge contrast then this would predict the change in strength of brightness induction that we measured. Perhaps the visual system does adjust for this bias partially, but cannot remove it completely.

One issue with this explanation is that we found changes in induction over a whole second, which appears to be longer than the temporal dynamics that Albrecht et al. (2002) measured. This can be explained, however, if you assume that the visual system integrates over the entire period of stimulus visibility. When the stimulus is presented only briefly, the visual system must calculate the strength of induction based on this elevated rate of firing, whereas longer presentation times can also integrate the lower,

sustained response. Lengthening the presentation time of the stimulus increases the potential contribution of the sustained signal relative to the transient signal, which would explain why we saw a graded decrease in effect as we increased *OnTime* all the way up to 1 s.

An important question is why Rossi and Paradiso's technique caused brightness induction to disappear, whereas ours did not. One possibility is that brightness induction does continue to occur, even in the 4 Hz sinusoidal modulation condition, but that the visual system averages over similar percepts when they occur over such short timescales. Perhaps when the modulation in brightness is small enough and fast enough, the visual system treats the modulation as noise, in which case the average would be a better representation of the true brightness. In our paradigm the mask is quite different from the inducing grating, greatly reducing the likelihood that the difference between the two stimuli is just due to internal noise, and thus indicating that the two percepts should not be averaged.

Given our results, one might predict that Rossi and Paradiso's data should have shown an increase in the strength of induction with modulation rate. If transient responses play a role, however, one would not expect their paradigm to show the same effect as our experiment. In their paradigm the modulating stimulus is continuously visible until the end of the trial, so any effect of the initial transient would be minimal. In addition, because the modulating stimulus was visible for many seconds, there is probably some visual adaptation to the brightness modulation. Thus, the adaptation to the real brightness modulation of the inducing region may additionally reduce the visual system's sensitivity to the weaker, induced brightness modulation. While this is speculative, it is clear that adaptation can have a big effect on stimulus visibility. When we piloted Experiment 3, we found that adapting to a visual mask of the same spatial frequency as the inducing stimulus made the brief presentation of the induction grating invisible. Further research using the Rossi and Paradiso paradigm should carefully investigate the potential role of adaptation.

Finally, there is a relatively high-level explanation of these results. The visual system is probably adapted to quickly extracting brightness from a brief glimpse (after all, this is the task it must solve every day, between saccades). It is not surprising, however, that the shorter the glimpse the less accurate the perceived brightness. Meanwhile, it is less clearly advantageous to respond to the situation where a constantly visible stimulus is surrounded by a rapidly flickering surround. This latter situation is somewhat similar to the visual experience during smooth pursuit eye movements when an object moves over a variable luminance background. Perhaps the visual system is adapted to down-weight such modulations. Indeed, as discussed earlier, the visual system is known to respond strongly to transient signals, and to quickly adapt to constant signals.

6.2. Filling-in

Our results do not suggest that the speed of filling-in plays a significant role in the temporal limits of brightness perception, at least for the *OnTimes* we tested. In particular, our results are not compatible with the slower estimates of filling-in, such as the 9–14°/s that Davey et al. (1998) estimated would explain Rossi and Paradiso's first experiment or Davey et al.'s estimate of 11–29°/s, for their own experiments on the COC illusion. Our results could, however, be compatible with significantly faster filling-in speeds, presuming they were fast enough that there should be no difference between the two stripe widths we tested at 58 ms. Thus our results are compatible with Rossi and Paradiso's estimate of 140–180°/s, based on measuring the apparent phase lag in the induced luminance in their 3rd experiment. At that speed, the largest stripe width we tested (10.6°) would be filled-in after only 29–37 ms, assuming that there are no other temporal delays other than calculating filling-in.

Recently, recordings in cat areas V1 and V2 suggest that filling-in is actually much faster than 140–180°/s (Hung, Ramsden, & Roe, 2007). Based on temporal correlation of spikes from pairs of neurons, they estimated that filling-in within V1 travels between 1300–2400°/s, and as fast as 4000°/s from V1 to V2. While there are likely many differences between cat V1 and human V1, these data do suggest 180°/s is likely too conservative an estimate of the speed of filling-in. On the other hand, Huang and Paradiso (2008) have found evidence of much slower filling-in in monkey V1. They found many cells fired much earlier to a contrast border than to the interior of a large uniform region. Based on these results the authors calculated that filling-in travels at about 270°/s. The authors further speculate that after adjusting for differences between human and monkey V1, their results would be compatible with speeds of 150–225°/s in humans. The wide difference between these two experiments may be due to the different measure used (spike-timing correlation, vs. change in mean firing rate). In any case, both experiments, as well as our own results, suggest that the 9–14°/s rate of filling-in necessary to explain Rossi and Paradiso's first experiment is improbably slow.

7. Conclusions

Our work suggests the generation of an initial brightness percept can occur very quickly and that the perceived brightness of a region depends on the surrounding context even for very short presentations. Both brightness induction and White's illusion were visible at the shortest times we tested, suggesting that both of these illusions are generated very quickly in the visual system. We did find, however, that the initial brightness percept can change if more processing time is allowed, particularly for brightness induction. In contrast to previous experiments, there was no indication that filling-in plays an important role in the temporal dynamics of brightness perception, but this may be because filling-in is too fast to significantly limit the speed of brightness perception.

Acknowledgments

We thank Paul Hammon, Don MacLeod, Stuart Anstis, and Karen Dobkins for helpful discussions about this work. This work was supported by the National Science Foundation under NSF Career Grant No. 0133996 to V.R. de Sa. A.E. Robinson was supported by NSF IGERT Grant #DGE-0333451 to G.W. Cottrell/V.R. de Sa.

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.

Anderson, B. (1997). A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions. *Perception, 26*, 419–431.

Blakeslee, B., & McCourt, M. E. (2008). Nearly instantaneous brightness induction. *Journal of Vision, 8*(2):15, 1–8. Available from http://journalofvision.org/8/2/15/.

Blakeslee, B., & McCourt, M. E. (1999). A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Research, 39*, 4361–4377.

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

Dakin, S. C., & Bex, P. J. (2003). Natural image statistics mediate brightness 'filling-in. *Proceedings of the Royal Society of London, 270*, 2341–2348.

Davey, M. P., Maddess, T., & Srinivasan, M. V. (1998). The spatiotemporal properties of the Craik-O'Brien-Cornsweet effect are consistent with "filling-in". *Vision Research, 38*, 2037–2046.

De Valois, R. L., Webster, M. A., De Valois, K. K., & Lingelbach, B. (1986). Properties of brightness and color induction. *Vision Research, 26*, 887–897.

Devinck, F., Hansen, T., & Gegenfurtner, K. R. (2007). Temporal properties of the chromatic and achromatic Craik-O'Brien-Cornsweet effect. *Vision Research, 47*, 3385–3393.

Enns, J. T., & Di Lollo, V. D. (2000). What's new in visual masking? *Trends in Cognitive Sciences, 4*, 345–352.

Glad, A., & Magnussen, S. (1972). Darkness enhancement in intermittent light: An experimental demonstration. *Vision Research, 12*, 1111–1115.

Gunther, K. L., & Dobkins, K. R. (2005). Induction effects for heterochromatic flicker photometry (HFP), heterochromatic brightness matching (HBM), and minimally distinct border (MDB): Implications for the neural mechanisms underlying induction. *Journal of the Optical Society of America, 22*, 2182–2196.

Huang, X., & Paradiso, M. A. (2008). V1 response timing and surface filling-in. *Journal of Neurophysiology, 100*, 539–547.

Hung, C. P., Ramsden, R. M., & Roe, A. W. (2007). A functional circuitry for edge-induced brightness perception. *Nature Neuroscience, 10*, 1185–1190.

Kingdom, F., & Moulden, B. (1988). Border effects on brightness: A review of findings, models and issues. *Spatial Vision, 3*, 225–262.

Magnussen, S., & Glad, A. (1975). Temporal frequency characteristic of spatial interaction in human vision. *Experimental Brain Research, 23*, 519–528.

McCourt, M. E., & Foxe, J. J. (2004). Brightening prospects for early cortical coding of perceived luminance: A high-density electrical mapping study. *Neuroreport, 15*, 49–56.

Paradiso, M. A., & Hahn, S. (1996). Filling-in percepts produced by luminance modulation. *Vision Research, 36*, 2657–2663.

Paradiso, M. A., & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research, 31*, 1221–1236.

Pelli, D. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Reeves, A. (2007). An analysis of visual masking, with a defense of stopped processing. *Advances in Cognitive Psychology, 3*, 57–65.

Robinson, A. E., Hammon, P. S., & de Sa, V. R. (2007). Explaining brightness illusions using spatial filtering and local response normalization. *Vision Research, 47*, 1631–1644.

Rossi, A. F., & Paradiso, M. A. (1996). Temporal limits of brightness induction and mechanisms of brightness perception. *Vision Research, 36*, 1391–1398.

Rossi, A. F., & Paradiso, M. A. (1999). Neural correlates of perceived brightness in the retina, LGN, and striate cortex. *Journal of Neuroscience, 19*, 6145–6156.

Schroeder, C. E., Mehta, A. D., & Givre, S. J. (1998). A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cerebral Cortex, 8*, 575–592.

Todorovic, D. (1997). Lightness and junctions. *Perception, 26*, 379–394.

White, M. (1979). A new effect of pattern on perceived lightness. *Perception, 8*, 413–416.

Chapter 2, in full, is a reprint of the material as it appears in Robinson, AE, & de Sa, VR. (2008). Brief presentations reveal the temporal dynamics of brightness induction and White's illusion. *Vision Research*, 48, 2370-2381. The dissertation author was the primary investigator and author of this paper.

Alan E. Robinson & Virginia R. de Sa

Department of Cognitive Science
University of California, San Diego

Dynamic brightness induction causes flicker adaptation, but only along the edges: evidence against the neural filling-in of brightness

Keywords: Dynamic brightness induction; flicker adaptation; filling-in; brightness perception

Please address correspondence to:

Alan Robinson
Department of Cognitive Science
University of California, San Diego
9500 Gilman Drive
La Jolla, California 92093-0515
USA

Phone: (858) 822 2421
FAX: (858) 534 1128
robinson@cogsci.ucsd.edu

Abstract

Is brightness represented in a point-for-point neural map that is filled-in from the response of small, contrast sensitive edge detector cells? We tested for the presence of this filled-in map using flicker adaptation. Subjects viewed illusory flicker caused by a dynamic brightness induction stimulus, with a modulating surround and a constant center. Thereafter flicker sensitivity was reduced when our test region was the same size as the constant center, but not for smaller, inset regions. This suggests that brightness induction does adapt cells along the contrast edge, but that there is no filled-in population of brightness selective cells to adapt.

Introduction

The perceived brightness of a surface depends on the brightness of the surfaces that surround it. This is known as brightness induction. One particularly well-known example of brightness induction is brightness contrast, in which the brightness of the surrounding surfaces *induces* a shift in the brightness of the center surface such that the apparent contrast is increased. Thus, a medium-gray square on a light background appears darker than a square with the same luminance on a dark background (Fig. 1a). Since the gray squares are physically identical (Fig. 1b), the response of photoreceptors in the retina must be the same. Since the appearance of the squares differ, however, there must be an area in the brain where that difference is represented. Even before information leaves the retina it is recoded in terms of contrast, rather than point-for-point luminance. While this would predict that the square's edges would appear different as a function of

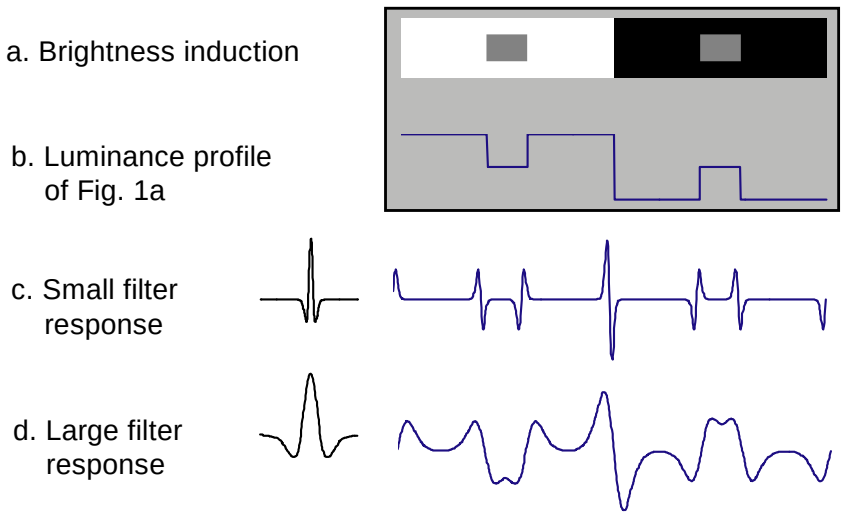


Fig. 1: Brightness induction and spatial filtering

their surround (Fig. 1c), it does not necessarily explain why the whole square appears to change, because the size of early receptive fields are much too small for cells responding to the center of the square to be influenced by the surround (induction can be clearly seen for figures where the induced region has an extent of 14°(Cornelissen, et. al, 2006)).

There are three commonly suggested explanations for how the change in appearance of the whole induced region is represented. The first is the **large receptive field** model: from the response of small receptive fields of early visual areas, larger, contrast sensitive receptive fields are constructed in later areas. If these later receptive fields are large enough to encompass the entire induced region and part of the surround, then their response would depend on both the luminance of the center and surround (Fig. 1d). The highly successful ODOG (Blakeslee & McCourt, 1999) and FLODOG (Robinson, Hammon, & de Sa, 2007) brightness perception models make use of such large-scale contrast sensitive filters, though they are agnostic as to how (or where) those large-scale filters are implemented.

A second possibility is known as **neural filling-in**. According to this theory at some post-retinal stage in the visual system there exists a point-for-point, retinotopic map of the brightness across the entire visual field. This point-for-point representation is generated by analyzing all the contrast edges, as determined by relatively small contrast-sensitive filters, determining the brightness on either side of those edges, and then finally propagating that brightness via a spread of activation from edges outward until the whole point-for-point representation is complete. One of the best known versions of this theory is Grossberg ‘s FACADE model (e.g., Grossberg & Todorovic, 1988), which can account

for a range of brightness illusions, including brightness induction. The filling-in model differs from the large receptive field model in several ways. Consider a small, uniformly colored rectangle which grows in size. As it gets bigger the amount of neurons which represent it in the filled-in map also grows. Meanwhile, the neurons which represented its filled-in brightness when it was smaller now just represent the brightness of the interior of the rectangle. In the large receptive field model, however, neurons respond to contrast at different scales. So, as the rectangle grows, the population that represents it shifts from neurons which respond to small spatial frequencies to those selective for large spatial frequencies.

The third theory suggests there is no point-for-point filling-in, or large contrast-sensitive receptive fields. Rather, the local edge responses themselves represent the whole brightness of the induced region, and the visual system directly infers from this signal the overall brightness of each region. This has been called the **symbolic filling-in** theory, since the visual system infers the appearance of the induced square based on the low-level input, but does not then re-represent that appearance in an 'analog', or point-for-point retinotopic map. We are unaware of any concrete model of how this would be implemented in the brain, but the general idea has been suggested by several authors (e.g. Dennett, 1992).

Most of the experiments conducted on this issue have focused on whether there is a retinotopic, point for point filled-in map of brightness. Rossi & Paradiso (1996) have argued that temporal limits on the perception of brightness induction reflect a point-for-point representation which takes a long time to fill-in. Their experiments showed that

temporally modulating the surround only induced a change in the brightness of a center region when the modulation rate is relatively slow (< 5 Hz). Furthermore, for larger induced regions, the temporal modulation rate that causes visible induction is even lower, dropping to below 2Hz for sufficiently large induced regions. This is consistent with the neural filling-in theory, where the propagation, or filling-in, of brightness takes time which is dependent on the distance that the signal must travel from contrast edges. While this result is certainly compatible with filling-in, it may instead be due to some yet-as-unknown factor, as two recent studies (discussed below) have shown that induction is visible at much higher speeds than Rossi & Paradiso's 1996 results suggest.

Blakeslee & McCourt (2008) investigated the timecourse of brightness induction using the grating induction illusion. Subjects viewed a sinusoidal grating that was modulated in counterphase. This induced a modulating grating 180° out of phase in a medium-gray test stripe that bisected the modulating grating. They found little to no change in phase between the induced grating and the inducing grating as the modulation rate increased from 2Hz to 24Hz. This suggests that induction sufficient to drive the motion percept occurs at the same speed as the perception of the inducing grating. While this is not particularly compatible with Rossi & Paradiso's theory of slow filling-in, it is compatible with filling-in, so long as filling-in is almost instantaneous.

We (Robinson & de Sa, 2008) have also shown that brightness induction appears to happen much quicker than suggested by Rossi and Paradiso, even when using induction stimuli closely modeled on Rossi and Paradiso's experiments. Instead of measuring induction strength while the surround was constantly modulated, however, we

used brief, static presentations, followed by a noise mask. We found that induction was clearly visible and quite strong, even when the brief presentation was as short as 58ms, and there was no indication that the spatial scale of the induced region had any influence on the temporal limits of induction. These results differ from Rossi & Paradiso's data across several dimensions, and suggest that the temporal properties measured by Rossi & Paradiso are actually due to some other factor, not yet identified. Note that our results, however, are still compatible with filling-in, so long as it happens on a sufficiently fast time-scale.

Some psychophysics paradigms unrelated to brightness induction also support the hypothesis that brightness spreads from contrast edges at a relatively slow speed. Paradiso & Nakayama (1991) studied brightness percepts elicited by briefly presenting a large white disk and then masking it with a smaller pattern, such as a black circle with a white outline. At short inter-stimulus-intervals the mask made the white disk appear as though it had a dark hole in the center, suggesting that the addition of a contrast edge blocked the completion of a slow filling-in process.

Paradiso & Hahn (1996) showed that steadily decreasing or increasing the luminance of a disk led to a slightly delayed change in the perceived brightness at the center of the disk. This lag is suggestive of a relatively slow filling-in process, though it is also compatible with the theory that the response of larger receptive fields in later visual areas are built from the response of smaller receptive fields in early visual areas. So long as people are conscious of the response in both early and later areas, the later areas would necessarily lag in time.

fMRI experiments in humans have produced mixed evidence for the filling-in of brightness information from contrast edges. Cornelissen, et al, (2006) measured the retinotopic response to viewing a static 14° wide gray disk while the surround modulated in brightness. In both V1 and V2 they found a very strong response in the voxels corresponding to the retinotopic location of the contrast edge between the center and surround. As the retinotopic distance from the edge increased, the BOLD signal decreased, with a similar fall-off for the regions corresponding to the center and the surround. This fall-off was quite slow, suggesting a very spatially broad response to edges; broader than would be expected due to the blurring caused by the BOLD point-spread function. While this could be the signature of filling-in, they argue that it is actually only evidence of a spatially diffuse response to edges. This argument is supported by their findings that a similar broad edge response was found when subjects viewed a modulating checkerboard background (which induced no change in brightness of the center, static region). Furthermore, a broad edge response was found in the surround when the center disk modulated in luminance and the surround was held constant, even though in this situation the surround appears constant in brightness. Other studies that purport to have found fMRI evidence for filling-in may well have been measuring this broad edge response instead (e.g. Boyaci, et al, 2007).

Pereverzeva & Murray (2008) followed up on Cornelissen's experiment but with an interesting twist – they tested induction at different luminance's for the central, static disk. When the luminance was low, little illusory change was seen in the central disk from modulating the surround. When the central disk had a higher luminance, however,

the brightness illusion was clear and quite strong. They found the largest V1 BOLD modulation in this latter condition, suggesting that the BOLD signal corresponds at least in part to the perceived brightness. Like Cornelissen they measured BOLD modulation as a function of distance from the contrast edge, and found a similar falloff in signal change as the distance from the edge increased. Puzzlingly, however, this was true even in the low-luminance condition, where no brightness induction was seen, again raising the question of whether the BOLD signal corresponds to perceived brightness as the authors suggest, or just a broad edge response, as Cornelissen argues. In addition, it is difficult to eliminate the possibility that broad edge response found by both studies is actually an fMRI artifact.

Single unit recordings have produced some evidence for filling-in, but at radically different speeds. Recordings in cat areas V1 and V2 suggest that a signal does propagate from contrast edges into the center of uniform regions (Hung, Ramsden, & Roe, 2007). Based on temporal correlation of spikes from pairs of neurons, they estimated that filling-in within V1 travels between 1300-2400°/s, and as fast as 4000°/s from V1 to V2, although aspects of their data led the authors to suggest the spreading activation in V1 was unrelated to perceived brightness. In contrast, Huang & Paradiso (2008) have found evidence of much slower filling-in in monkey V1. They found many cells fired much earlier to a contrast border than to the interior of a large uniform region. Based on these results the authors calculated that filling-in travels at about 270°/s. The wide difference between these two experiments may be due to the different measure used (spike-timing correlation, vs. change in mean firing rate), or the difference in species.

In the present work we introduce a new psychophysical paradigm for measuring the presence of neural filling-in, based on flicker adaptation. Prolonged viewing of a high-contrast flickering spot causes reduced sensitivity to that flicker, and in particular, elevated contrast thresholds for low-contrast flicker (see Schietin and Spillman, 1987, for a brief review). This reduction in sensitivity may be due to neural fatigue of the cells which represent the flickering region, or to a process which calibrates sensitivity in order to reduce the firing rate to frequently encountered stimuli (the precise explanation does not matter for the purpose of our research). In our current study we explore if flicker adaptation is found for regions where illusory flicker is caused by brightness induction from actual flicker in the surround. We will show that induction does cause strong flicker adaptation, suggesting that it isolates a population of neurons which represent the change in appearance of the induced region. This reduction is strongest when the inner edges of the flickering inducer are aligned with the outer edges of the test region. We will also show that the strength of flicker adaptation is drastically reduced when the test region is shrunk (that is, inset) a few degrees, suggesting that the population of cells that are adapted are primarily or even exclusively along the contrast edges, and if there is a filled-in population that represents the brightness of the induced region, it is not susceptible to adaptation.

Experiment 1: Adapt to induction, detect polarity

In this experiment we measure the strength of induction-derived flicker adaptation for two conditions: when the test region is exactly *aligned* with the inner edge of the

inducer, and where the test region is significantly *inset* relative to the edges of the inducer (Fig. 3). In the aligned condition two forms of adaptation are possible: first, we expect some edge adaptation because the contrast modulation at the border between the inducer and the inducing region exactly aligns with the contrast modulation between the test region and the background. Second, if there is a neural filling-in of brightness then we would expect those neurons to be adapted across the entire region where the induced flicker is seen. In the inset condition the edges of the inducer and the test region do not overlap, so the only type of adaptation that can remain is to the posited filled-in brightness modulation. Thus, if there is neural filling-in we should expect to measure an effect of adaptation in both the aligned and inset conditions.

Methods

Subjects

Two moderately psychophysically experienced subjects participated. Both were naïve to the purpose of the experiment.

Apparatus

Stimuli were presented on a 21" NEC FE2111SB CRT driven by an ATI RADEON 7000 VE video card at a refresh rate of 85Hz. Display luminance was linearized using a 256-entry color lookup table that drove a 10-bit DAC. A Cambridge Research Systems ColorCal colorimeter was used to select the appropriate lookup table values. The lookup table was apportioned so that the full resolution of the DAC was available for luminance values between 37 cd/m^2 and 65 cd/m^2 (in steps of 0.15 cd/m^2). The remaining entries covered the rest of the 0 to 102 cd/m^2 range in much coarser

increments. This allowed us to adapt subjects with maximum contrast inducers, while still allowing us to also generate relatively low contrast test stimuli.

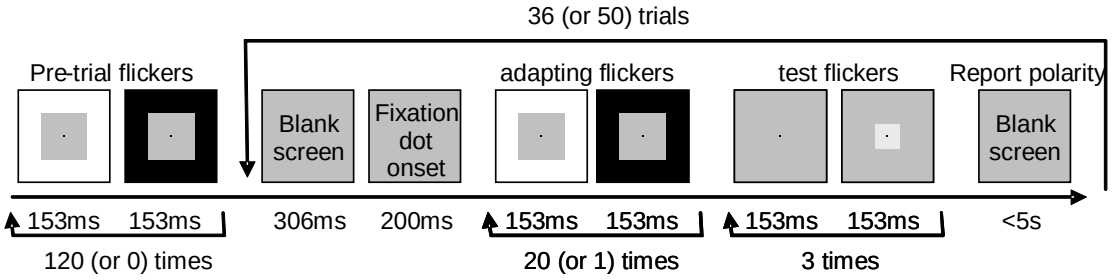


Fig. 2: Experiment

A chinrest was used to maintain a viewing distance of 72cm. Stimuli were generated and displayed using Matlab running the Psychophysics Toolbox, version 2.54 (Brainard, 1997; Pelli, 1997) on a Windows XP computer. The experiment was run in a dark room and subjects adapted to the light level for 3 minutes before collecting data. The same apparatus was used in all experiments.

Stimuli and procedure

Our paradigm is outlined in Fig. 2 Each large rectangle represents a single screen shown to the subject, with the duration that screen was visible listed below it. The experiment started with the screen on the far left, and proceeded rightward. The arrows show loops in the time line, causing groups of screens to be repeated multiple times (such as the pre-trial flickers, which were shown a total of 120 times in the *adapt* trials, or 0 times in the *no-adapt* trials.)

Contrast thresholds were measured in two different kinds of sessions: *adapt* and *no-adapt*. In the adapt session subjects first view a fixation dot centered in a gray rectangle (50.62 cd/m^2) surrounded by a larger rectangle whose luminance is modulated by a 3.3Hz square-wave (153ms per frame) between 0 and 102 cd/m^2. The luminance change of the surround induces a change in the apparent brightness of the physically static center. Subjects adapt to 120 cycles of flicker. This *Pre-trial* flicker is meant to build up a high level of adaptation before any measurements are made. Subjects are instructed to fixate the dot at the center of the screen during this, and all subsequent parts of the experiment. After the pre-trial flickers the measurement trials start.

A trial in an adapt session consists of the following screens. First, a gray screen (50.62cd/m^2) is shown for 306ms, serving as a very brief inter-trial break, and then a fixation dot is drawn, which remains on the screen for the rest of the trial. After an additional 200ms delay subjects view 20 full-contrast *adapting flickers*, which are identical to the pre-trial flickers in all ways except in total number. These flickers serve to keep subjects in an adapted state. Then three *test flickers* are shown. The test flickers consist of a gray screen for 153ms, followed by a gray rectangle that is either a small increment or small decrement relative to the background luminance, for 153ms. On any given trial the test flickers are all either increments or decrements, and the subject's task is to report this using the keyboard after the 3rd flicker is shown (we will refer to this as the *polarity task*). The fixation dot disappears after the 3rd flicker to indicate a response is required. Subjects are encouraged to respond quickly but are given up to 5 s to respond before the trial times out. After the subject responds a new trial begins.

To measure the effect of adaptation we compared contrast thresholds for the adapt sessions to *no-adapt* sessions. In the no-adapt sessions no pre-trial flickers are shown, and just 1 “adapting flicker” is shown instead of 20 adapting flickers. We included one flicker to make sure that any masking effects, such as meta-contrast masking, were present in both the adapt and no-adapt sessions. In the no-adapt sessions we collected 50 trials per session; in the adapt sessions we collected just 36 (each trial took much longer to complete and we did not want to overly fatigue subjects).

To measure contrast thresholds we adjust the strength of the test flickers using a variable stepsize staircase. The staircase stepsize starts at 0.75cd/m^2 (1.5% Weber contrast) and is reduced by 10% every time the subject answers incorrectly on a trial, but is not allowed to decrease below the minimum resolution of the DAC. After correct trials the flicker strength is reduced by one stepsize; after incorrect trials it is increased by two stepsizes. If the staircase reaches zero flicker contrast the stepsize is also reduced by 10%

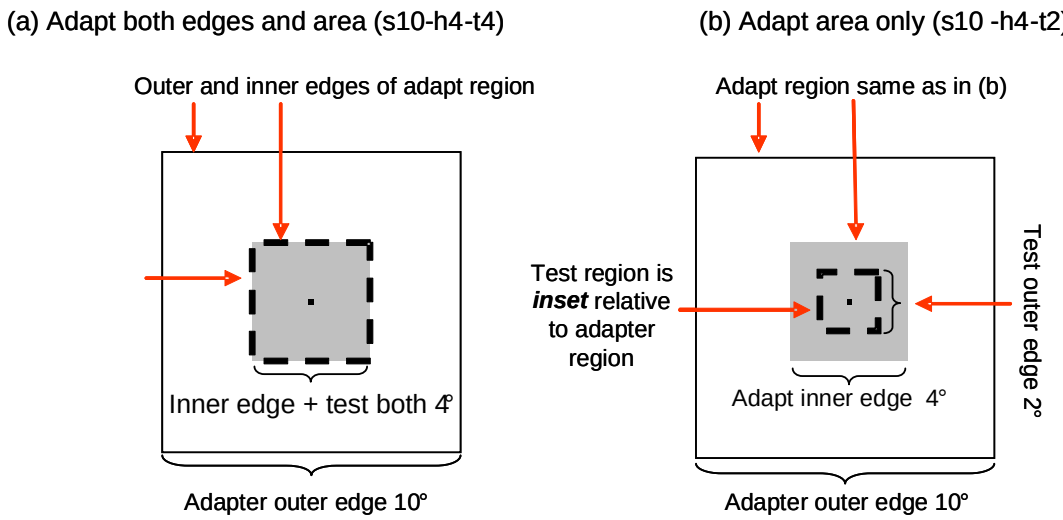


Fig. 3: aligned and inset conditions

and the staircase is reset to the minimum possible flicker greater than 0% contrast. After every 4 incorrect responses a single easy trial is introduced to keep the subjects' attention and motivation from decreasing due to frustration. The response on the easy trial, however, does not influence the location of the staircase on subsequent trials. These parameters were found to produce good coverage of the psychometric curve during pilot tests.

Pilot tests suggested that the adapt and no-adapt sessions produced rather different contrast thresholds. In order to minimize the number of non-informative trials, we started the staircases at different contrasts (in the no-adapt sessions 5.8% contrast, or ±2.94 cd/m^2 relative to the 50.62 cd/m^2 background; in the adapt sessions 13.3% contrast). These contrasts were significantly above threshold and subjects completed about 10 trials before failing to detect a significant number of test flickers. Pilot work suggested, however, that the initial contrast had little effect on the final measured contrast thresholds.

We tested 4 different configurations of inducing rectangles and test rectangles; 2 aligned (s10-h4-t4, and s14-h7-t7) and 2 inset (s10-h4-t2 and s14-h7-t2). These configurations are named for the width of the **S**urrounding inducer, the gray **H**ole in the inducer, and the **T**est rectangle (in degrees of visual angle). Two examples are sketched in Fig. 3 (not to scale; the dashed rectangles represent the size and location of the test flicker). For each configuration we ran an average of 8 sessions, split between adapt and no-adapt sessions.

Analysis

We fit a Weibull curve to the data for each condition to estimate the percent contrast necessary to elicit 75% correct on the polarity task. A bootstrap analysis using the psignifit package (Witchman & Hill, 2001b) was used to estimate the 95% confidence intervals of this threshold, using 9000 bootstrap tests per curve. We allowed the software to estimate up to a 5% lapse rate (Witchman & Hill, 2001a) when fitting the Weibull function to the data, to prevent occasional lapses in attention from significantly influencing the estimated 75% correct threshold

Subjects tended to make the occasional mistake on the first few trials of a session, somewhat independent of condition, so we discarded the first 3 trials of each session to minimize the noise this would introduce in estimating the threshold.

Results

The percent contrast needed to reach a 75% correct performance threshold for each condition is shown in Fig. 4 (Error bars denote 95% confidence intervals in this, and all following figures. Any reduction in sensitivity due to adapting is indicated by the 'adapt' bars being further to the right than the no-adapt bars). To measure the decrease in sensitivity due to adaptation we compare this contrast threshold between adapt and no-adapt sessions. For the aligned conditions the difference is quite large for both subjects, showing that adapting to induced flicker can significantly reduce the visibility of low-contrast flickering regions. For the inset conditions, however, the difference between the adapt and no-adapt sessions is quite small, with no consistent trend across subjects, and is generally no larger than the 95% confidence intervals. This suggests that adapting to

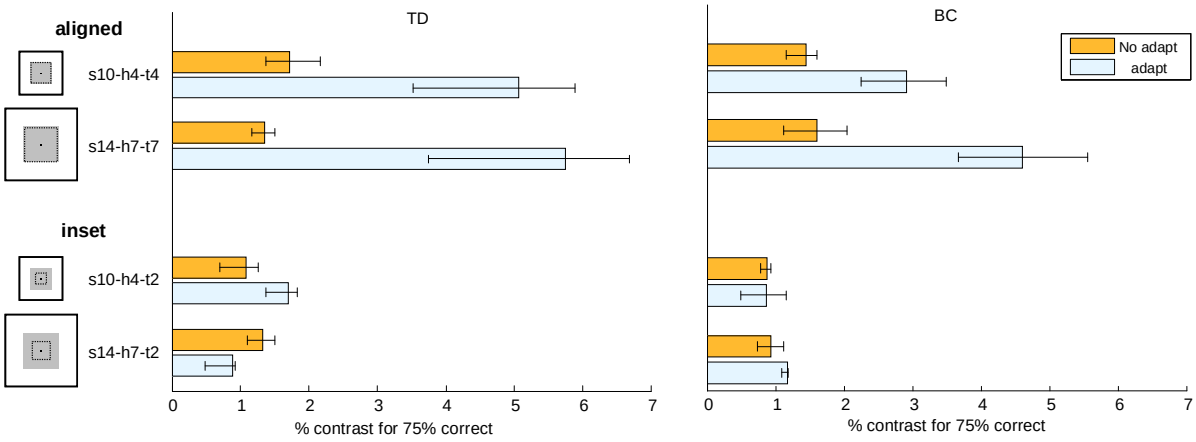


Fig. 4: Experiment 1: induced flicker adaptation

induced flicker does not reduce sensitivity to low-contrast flicker when the edges of the test region are inset from those of the inducer.

The results of Experiment 1 suggest that induced flicker adaptation is primarily driven by adaptation of cells near the contrast edges. We found no evidence that sensitivity is also reduced in the areas removed from the edges, even though flicker of the inducing rectangle caused an apparent flicker in the brightness of the entire gray central rectangle. This seems at odds with the theory that brightness information is neurally filled-in from contrast edges, since if there is a filled-in population we would expect those neurons to also show adaptation and therefore reduce flicker sensitivity across the entire induced flicker region.

Experiment 2: Adapt to induction, detect cycle count

In Experiment 1 we found that edges played a primary role in induction-derived flicker adaptation. In Experiment 2 we attempt to replicate this finding using a different

task. This was motivated by the informal observation during Experiment 1 that the general presence of flicker could be detected on some trials even when the polarity of that flicker was unclear. Thus, we wanted to determine if there is something special about polarity judgments. In our new task subjects counted the number of cycles of test flicker on each trial. We also changed the scale of the adapting and testing regions to rule out the alternate hypothesis that small test regions are not affected by adaptation (all of the inset test regions were only 2° wide in Experiment 1).

Methods

Two psychophysically experienced subjects who were naïve to the purpose of the experiment participated. One author (AR) also participated.

The paradigm remains similar to that sketched in Fig. 2 , and is identical to that used in Experiment 1 except as discussed in this section.

The major change was that the subject's task was to count the number of cycles of test flicker at the end of each trial. On trials where only one or two flickers were shown the remaining flickers were "drawn" at zero amplitude, insuring that total trial timing was the same irrespective of flicker count. The fixation dot always disappeared 6x153=918ms after the final adapt flicker so that trial length would not serve as a cue to the number of flickers.

We implemented two other small changes to ensure that the difference between no-adapt and adapt sessions did not depend on extraneous factors. First, we were concerned that performance on adapt trials might have been lowered because there was

no cue immediately proceeding the test flickers, and thus performance could decrease if subjects' attention wandered during adaptation. Therefore, we had the fixation dot change on the final cycle of the adapt flickers from white to black. In contrast, no-adapt sessions had only 1 'adapt' flicker instead of 20, so there was little reason to worry that attention might drift. Note, however, that we implemented the fixation dot change in both adapt and no-adapt sessions to keep the stimulus properties equal.

The other issue we addressed was that in the no-adapt trials there is only 306ms + 200ms = 506ms of delay after the subject's response (see Fig. 2). Since the no-adapt trials had a single 'adapt' flicker (to equalize any masking, as discussed in Experiment 1 methods), it is possible that a small amount of adaptation might build up over the course of the 50 trials. Thus, in the no-adapt sessions we introduced an additional delay before the single adapt flicker showing the fixation dot for 1.3s instead of just 200ms. We did not, however, increase the delay in the adapt sessions.

We elected to collect 50 trials in both the adapt and no-adapt sessions, since in Experiment 1 subjects had not complained of any fatigue. The adapt and no-adapt sessions were also made more similar by using the same staircase parameters for both session types (in particular, both staircases started at 13.3% Webber contrast).

We tested 4 different configurations of inducing rectangles and test rectangles; 2 aligned (s10-h4-t4, and s15-h8-t8) and 2 inset (s10-h4-t2 and s15-h8-t4). For each configuration we ran an average of 9 sessions, split between adapt and no-adapt sessions.

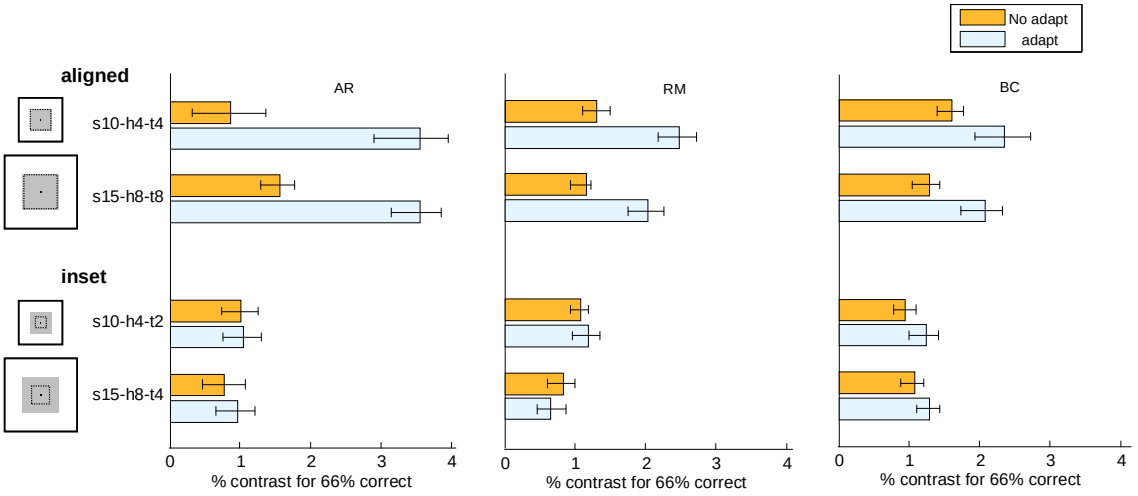


Fig. 5: Experiment 2: count the number of test flickers (3-AFC)

We determined the contrast threshold necessary for 66% correct performance (halfway between chance and perfect performance for a 3-alternative forced choice response), using the same psignifit package settings as in Experiment 1.

Results

The results are shown in Fig. 5. All subjects show much higher contrast thresholds after adaptation in the aligned condition. The effect is somewhat weaker for subject BC, and his data are somewhat more noisy, especially in the s10-h4-t4 condition, such that the 95% confidence intervals almost overlap. This could plausibly be due to extraneous eye movements, since the receptive fields near the fovea are smaller than those in the periphery, which would mean that small eye movements would be more detrimental to edge-based adaptation. This hypothesis is in agreement with the observation that all subjects show wider 95% confidence intervals for the s10-h4-t4 adapt sessions than the s15-h8-h8 adapt sessions.

In contrast to the aligned condition, the data from the inset conditions show no consistent evidence of adaptation, with overlapping 95% confidence intervals for all of the subjects, for both the s10-h4-t2 and s15-h8-t4 configurations.

These results suggest that induced flicker adaptation is confined to edges. We found no clear evidence of adaptation to areas interior to the edges. This replicates the findings from the first experiment, and shows that this result is not limited to the polarity task.

Experiment 3: Jittering edges and slower flicker

Experiments 1 and 2 suggest that induced flicker adaptation acts solely on cells along the contrast edge. This suggests that there is no pool of filled-in neurons to adapt that represent the brightness of the area removed from the edges. What if, however, there is such a pool of filled-in neurons, but our flickering inducer fails to stimulate them sufficiently to cause a measurable amount of adaptation? In Experiment 3 we address two possible reasons this might occur.

First, filling-in is thought to be a two-stage process: initially edge units respond to local contrast, and then this information is propagated to units representing uniform areas. If, however, the edge units themselves become sufficiently adapted then they might be less likely to significantly modulate the response of the filled-in neurons. This hypothesis would be particularly compelling if our flickering adapter did not cause a clear brightness change across the whole induced region. Subjectively, however, the brightness of the induced region did modulate, weakening this hypothesis. Even so adaptation in the edge-selective neurons might reduce the amount of adaptation in the filled-in neurons, such

that we could no longer detect it reliably. To address this, we randomly modulated the location of the inner and outer edges of the inducing rectangle, so that there would be significantly less adaptation to local edges, while still modulating the supposedly filled-in neurons.

Second, it is possible that the 3.3Hz modulation rate of the inducer was just too fast for filling-in to complete. If filling-in did not complete, one would expect significantly less adaptation, particularly in the center of the filled-in region. De Valois, et. al, (1986) found that the strength of induction does drop precipitously between 2 and 4 Hz, at least in some subjects. Note, however, that for 5 out of 6 subjects, induction did not disappear completely even at 8Hz modulation rates. If the weakening of induction is related to the slow speed of filling in, as argued by Rossi and Paradiso (1996), then this could explain why we could not measure any adaptation for the filled-in cells. Note that the slow speed of induction (and therefore filling-in) is under some dispute, however: Blakeslee & McCourt (2008) have suggested that induction is nearly instantaneous, and we (Robinson & de Sa, 2008) have argued that a single 58ms exposure is enough time for induction to occur (at 3.3hz each frame of the inducer is shown for 153ms).

Even if filling-in is fast enough to complete for the 3.3Hz modulation rate that we used, it is true that prolonged viewing of modulation at this “high” speed would tend to reduce the apparent strength of induction, which could reduce the potential for adaptation. To this end we also explored if a ~1Hz (506ms per frame) modulating inducer would reveal greater adaptation. 1Hz is slow enough that even by the slowest estimates

filling-in should complete, and furthermore, 1Hz modulation causes stronger induction than 3.3Hz.

We also adjusted the no-adapt sessions to be equally fatiguing as the adapt sessions. If fatigue was higher in the adapt sessions this could potentially decrease contrast sensitivity in those sessions. The expected size of this effect is small, but since we were eliminating potential reasons why no effect was found in the inset conditions, we wanted to also eliminate potential biases in the opposite direction. The no-adapt sessions now take exactly as long as the adapt sessions, and have the same number of flickers, but the location of the flicker is moved so that it does not induce any change in the apparent brightness around fixation.

Finally, we used just one cycle of test flicker at the end of the trial. This protected against the possibility that by the time 3 test flickers had been shown the filled-in neurons had already un-adapted significantly. Since every trial had exactly one flicker we used the polarity task from Experiment 1.

Methods

We used the same three subjects as in Experiment 2.

Our paradigm differs from Experiment 1 in the following ways. First, after every full cycle (2 frames) of the pre-adapt or adapt flickers, the inner and outer edges of the inducing rectangle were moved within a 1° range (in step sizes of $1/32^{nd}$ of a degree). For the s13-h6-t2 configuration this meant that the inner edge of the inducing rectangle could be anything between 6° and 7 ° in width. The size of the outer inducing edge was always

expanded by the same amount as the inner edge, which loosely preserved the degrees of visual angle subtended by the inducing rectangle.

The no-adapt sessions were changed to have exactly the same temporal properties as the adapt sessions. Instead of changing the number of flickers between sessions, we changed the location of the flickers in the no-adapt sessions so that the flicker would contribute minimally to any visual adaptation at the target location. In no-adapt sessions all but 1 of the pre-adapt and test flickers were shown in a rectangle (4.15° by 32°) abutting the top edge of the screen. These *top-flickers* covered on average the same amount of visual angle as the flickering region centered around the test rectangle in the adapt sessions. They did not induce any apparent change in the brightness of the rest of the screen.

Thus, in no-adapt sessions, subjects first viewed 120 flickers occurring at the top of the screen while fixating at the center of the screen. Then, each trial was preceded by 19 top-flickers, followed by a single regular induced flicker surrounding the test patch, and then finally the test flicker was shown. In adapt sessions no top-flickers were shown; all adapting flickers were centered around the fixation point, as in Experiment 1.

As in Experiment 2 we had the fixation dot change on the final cycle of the adapt flickers from white to black to cue subjects that the test flickers were about to appear.

We collected data for 3.3Hz and 1Hz flicker speeds. Experimental sessions took dramatically longer in the 1Hz condition, so in order to maximize the amount of useful data we adjusted the staircase procedure to start just a little above the contrast needed for

100% performance. A brief pilot experiment determined that a single set of parameters achieved this for all of our subjects: 3% initial contrast, with an initial stepsize of 0.5% contrast. The same parameters were used in the 3.3Hz condition. We collected 50 trials per session in the 3.3Hz condition and 30 trials per session in the 1Hz condition. We collected an average of 4 sessions for each speed (split between adapt and no-adapt sessions).

We tested just one configuration of inducing rectangle and test rectangle: s13-h6-t2. We did not include an aligned configuration because the location of the inducing rectangles edges varied rapidly across the entire session, as described at the beginning of this section.

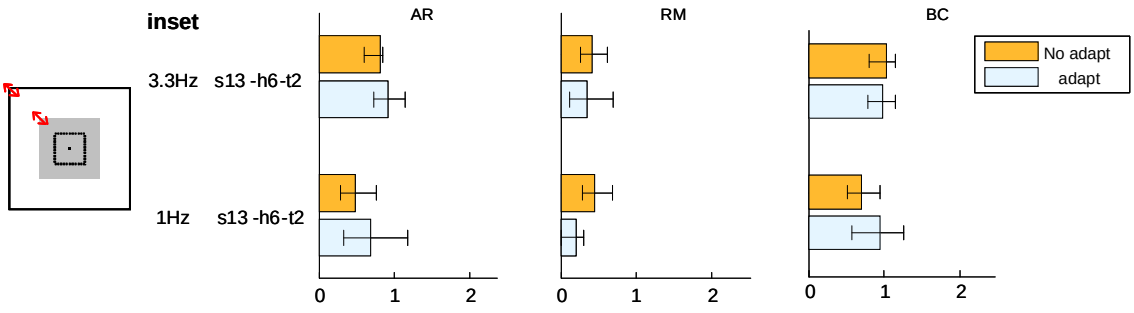


Fig. 3: Experiment 3: variable sized adapter, single flicker polarity test

Results

We determined the contrast threshold necessary for 75% performance, using the same psignifit package settings as in Experiment 1. The contrast thresholds are shown in Fig. 6 . In both the 1Hz and 3.3Hz conditions there was no evidence of increased contrast thresholds in the adapt condition. These data replicate the general finding from the first

two experiments that adapting to induced flicker does not cause flicker adaptation in the interior of the induced region. Thus, the lack of adaptation does not appear to be due to confounds from overly rapid flicker in the inducer, or because adaptation to edge contrast prevented adaptation from occurring in the supposedly filled-in neurons.

General Discussion

We have shown that adapting to induced flicker can cause significant reductions in contrast sensitivity to real flicker, but only in the situation where the edges of the inducer are aligned with the edges of the test region. When the test region is inset, it appears that there is no reduction in sensitivity.

These data highlight the important role of edges and contrast in brightness induction. It is not particularly consistent with the theory that brightness information 'fills-in' from edge selective neurons into a point-for-point representation of brightness across the entire visual field. In particular, if such a filled-in representation exists, we would expect that it too is susceptible to flicker adaptation, just like the cells near the contrast edge, and that this adaptation would have some effect on contrast sensitivity in regions inset from the edges. Note that our data do not argue that cells only respond to the contrast edge; there may be cells which fire to the interior of a region who's apparent brightness is modulated, as some single unit recording experiments have indicated (e.g. Rossi & Paradiso, 1999). Our results do suggest, however, that the firing of such cells does not contribute significantly to the representation of brightness, or at the very least that they do not contribute to detecting brightness flicker at the speeds we tested.

There are some caveats, however, to our conclusion that brightness does not fill-in. It is possible that brightness is represented in a filled-in population, but for some reason that population is not susceptible to flicker adaptation. This would be quite surprising, since adaptation seems to be a relatively ubiquitous phenomena, but of course it cannot be ruled out.

Another, more plausible objection is that even though we adapted the supposedly filled-in neurons, we did not adapt the edge-selective neurons which were internal to the filled-in region. Thus, in the inset condition, the un-adapted edge-selective neurons were able to signal for the edges of the inset flickering rectangle. While the adaptation of the filled-in neurons may have occurred, this edge signal could have been present with no reduction in strength, and perhaps this is the signal that subjects used to govern their response. It is impossible to rule this out, though we have two arguments for why it seems implausible. First, if filling-in is an important component of representing brightness, then we would expect that interfering with it would cause some reduction in sensitivity – otherwise why would the visual system need to fill-in at all? Second, if filling-in neurons were adapted, but edge neurons were not, then one might expect some rather strange percepts in the inset condition – such as flickering edges with no center, sort of like a wireframe rectangle. At least subjectively, we did not observe any such percepts.

What about other theories of filling-in besides the point-for-point model? Our results are quite compatible with the symbolic theory of filling-in. Adapting the edge-

selective cells which code for filled-in regions reduces contrast sensitivity to other filled-in regions that are represented by those same cells.

Our results are also compatible with the theory that there are large contrast sensitive receptive fields which are built out of the response of smaller receptive fields. Adapting to a flickering inducer will affect all of the receptive fields at the appropriate scales to detect the contrast across the edges, and also the much larger receptive fields which code for the contrast of the center relative to the surround. These larger receptive fields would be poorly suited for signaling the presence of a smaller test rectangle, since the test rectangle stimulates a much smaller part of their positive (or negative) centers. Meanwhile, an intermediate sized receptive field, whose center just covers the smaller test rectangle is ideally suited for signaling its presence, and would not be adapted from the flickering inducer at all.

Conclusions

Many studies have been published on the issue of how brightness is represented in early visual areas. While there is no consensus, our results provide converging evidence from a new paradigm that a point-for-point, filled-in representation is unlikely. While we doubt this debate will be resolved soon, we suggest that it would be productive to consider alternatives to the point-for-point filled-in theory, and that it is time to start designing experiments that can differentiate between the other two possible methods of representing brightness described in this paper.

Acknowledgements

We would like to thank Don MacLeod for his help designing these experiments, and Paul Hammon, Stuart Anstis, and Karen Dobkins for helpful discussions about this work in general. This work was supported by the National Science Foundation under NSF Career Grant No. 0133996 to VR de Sa. AE Robinson and was supported by NSF IGERT Grant #DGE-0333451 to GW Cottrell/VR de Sa.

References

Blakeslee, B., & McCourt, M.E. (1999). A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Research*, *39*, 4361-4377.

Blakeslee, B., & McCourt, M. E. (2008). Nearly instantaneous brightness induction. *Journal of Vision*, 8(2):15, 1-8, http://journalofvision.org/8/2/15/

Boyaci, H., Fang, F., Murray, S. O., & Kersten, D. (2007). Responses to Lightness Variations in Early Human Visual Cortex. *Current Biology, 17*, 989-993.

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

Cornelissen, F. W., Wade, A. R., Vladusich, T., Dougherty, R. F., & Wandell, B. A. (2006). No Functional Magnetic Resonance Imaging Evidence for Brightness and Color Filling-In In Early Human Visual Cortex. *Journal of Neuroscience, 26*, 3634-3641.

De Valois, R.L., Webster, M.A, De Valois, K.K., & Lingelbach, B. (1986). Temporal properties of brightness and color induction, *Vision Research* 26, 887–897.

Dennett, D. C. (1992). "Filling in" versus finding out: A ubiquitous confusion in cognitive science. In H. L. Pick Jr., van den Broek, Paulus Willem & D. C. Knill (Eds.), *Cognition: Conceptual and methodological issues.* (pp. 33-49)

Grossberg, S., & Todorovic, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics, 43*, 241-277

Huang, X., & Paradiso, M.A. (2008). V1 Response Timing and Surface Filling-In. *Journal of Neurophysiology*, 100, 539-547.

Hung , C.P., Ramsden, R.M., & Roe, A.W. (2007). A functional circuitry for edge-induced brightness perception. *Nature Neuroscience*, 10, 1185-1190.

Paradiso, M.A. & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research*, 31, 1221–1236.

Paradiso, M. A. & Hahn, S. (1996). Filling-in percepts produced by luminance modulation. *Vision Research*, 36, 2657-2663.

Pelli, D. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437-442.

Pereverzeva, M., & Murray, S. O. (2008). Neural activity in human V1 correlates with dynamic lightness induction. *Journal of Vision, 8*(15):8, 1-10.

Robinson, A.E., Hammon, P.S., & de Sa, V.R. (2007). Explaining brightness illusions using spatial filtering and local response normalization. *Vision Research*, 47, 1631-1644.

Robinson, A.E., & de Sa, V.R. (2008). Brief presentations reveal the temporal dynamics of brightness induction and White's illusion. *Vision Research*, 48, 2370-2381

Rossi, A.F., & Paradiso, M.A. (1996). Temporal limits of brightness induction and mechanisms of brightness perception, *Vision Research*, 36, 1391–1398.

Rossi, A.F., & Paradiso, M.A. (1999). Neural correlates of perceived brightness in the retina, LGN, and striate cortex. *Journal of Neuroscience*, 19, 6145-6156.

Schietin, S., & Spillman, L. (1987). Flicker adaptation in the peripheral retina. *Vision Research, 27*(2), 277-284.

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

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

Chapter 3, in full, has been submitted for publication and is currently under review by the journal Vision Research. The dissertation author was the primary investigator and author of this paper.