Experience and Perception

by

Nathan Witthoft

B.A., English
University of Pennsylvania, 1999

Submitted to the Department of Brain and Cognitive Sciences in Partial
Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Cognitive Science

at the
Massachusetts Institute of Technology
September 2007

Signature of Author:________________________________
Department of Brain and Cognitive Sciences
September 4, 2007

Certified by:________________________________
Lera Boroditsky
Professor of Psychology
Thesis Supervisor

Accepted by:________________________________
Matthew Wilson
Professor of Neurobiology
Chairman, Department Graduate Committee

Experience and Perception

by

Nathan Witthoft

Submitted to the Department of Brain and Cognitive Sciences on September 4th, 2007, in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Cognitive Science

Abstract

To what extent can experience shape perception? In what ways does perception vary across people or even within the same person at different times? This thesis presents three lines of research examining the role of experience on perception. The first section presents evidence from synesthesia suggesting that learning can influence letter-synesthesia pairings and that associative learning can affect relatively early visual processing. The second section examines the role of linguistic categorization in color judgments, finding that language can play an online role even in a relatively simple color discrimination task. The final section examines how perception adjusts over relatively short time scales using face adaptation. The adaptation experiments show that adaptation to faces can improve recognition performance on famous faces. The results further demonstrate that these effects can be obtained without extensive training and that contrary to proposals from experiments using face spaces, that identity based adaptation effects can be found on trajectories which do not pass through the average face.

Thesis Supervisor: Lera Boroditsky

Title: Assistant Professor of Psychology

Acknowledgements

First, none of the work in this thesis would have been possible without the input and support of my advisor Lera Boroditsky, who has had enough faith in me to let me explore my multitude of interests making grad school the most enjoyable period of my life, but also the good sense to reign me in and make sure I finish things I start, which has also made it the most productive. I'd like to thank Jonathan Winawer, who has been a constant collaborator, sounding board, good friend, and has (as far as I can tell) cheerfully tolerated my manias almost continuously for the last six years, and only occasionally retaliated by showing pictures of me in his talks. I'm also grateful for the guidance, good advice, and genial dispositions of my other MIT committee members Nancy Kanwisher, and Molly Potter and would like to acknowledge Aaron Seitz for agreeing to be on my committee and providing many useful comments on my thesis. Sharon Thompson-Schill gave me my start in the field when I was still a delinquent English major, and I would not have even made it to graduate school without her support.

Lisa Wu worked hard as an RA on the color project in this thesis among many other things, and was a window onto the MIT undergrad world which appears to consist of Anime, slash fiction, and samba de amigos. Perhaps most importantly, Denise Heintze at MIT and Priscilla Feke at Stanford both deserve great credit for keeping me clothed, fed, and insured, and for shepherding me through the paperwork maze. Bart Anderson and Roland Fleming, kept me sane and happy in early days at MIT, and moreover, made me realize that perception was the most interesting thing a person could study.

I've had the good fortune of working in two departments during graduate school, and I'd like to thank the many people who made me feel welcome in the Stanford psychology department. Michael Ramscar, Kalanit Grill-Spector and Brian Wandell who served as mentors, and showed by example, what it means to do interesting but rigorous science. The members of the Institute for the Resurrection of Mathematical Psychology, Dan Yarlett, Nick Davidenko, and Adam November, became friends and collaborators, though I'm starting to wonder if there is a really is any such institute. Last, I'd like to thank all the members Lera's lab, which has steadily grown over the years, especially Daniel Casasanto who was there at the beginning, and Caitlin Fausey, who thinks I'm normal.

This thesis is dedicated to Yuko and Zach, who have followed me everywhere I needed to go, and for the rest, which they know.

TABLE OF CONTENTS

Chapter 1: Introduction

Associative Learning and Synesthesia

Can associative learning affect even relatively early visual processes in some people? Many theorists have proposed that associative learning plays an important role in perception and recognition. For example, numerous experiments demonstrate the importance of associative learning to our ability to recognize objects and faces despite changes in the image such as viewpoint (Logothetis et al, 1995; Sinha & Poggio, 1996; Wallis & Bulthoff, 1999). Furthermore, introducing false correlations between images has been shown to 'break' object invariance (Cox et al, 2005). Other research has shown that associative learning can bias the perception of ambiguous figures (Haijiang et al, 2006, Wallach & Austin, 1954).

In the first series of experiments I examine the role of associative learning in synesthesia. Synesthesia is experienced by a small subset of the population who in response to some stimuli such as letters, perceive additional sensations which are not part of the stimulus such as color. Most synesthesia researchers have tended to the view that the relationship between the synesthetic trigger (often referred to as the inducer) and the synesthestic percept (often referred to as the concurrent) is shaped by innate alignments between sensory dimensions. This is somewhat counterintuitive since many of the most prevalent forms of synesthesia seem to involve learned sequences such as numbers, letters, days of the week, and months (Rich & Mattingley, 2002). The nativist account is bolstered in part by experimental evidence showing that normal subjects show regular patterns in cross-modal associations, such as judging that higher pitched sounds are brighter than lower pitched ones (Marks, 1975). Other researchers have argued that infants are synesthetic, with differentiation in sensory dimensions arising as a result of pruning connections between different sensory areas (Maurer, 1997). Based on this kind of evidence, researchers have proposed that synesthesia might result from a failure in some individuals to prune connections between sensory areas during early development (Ramachandran & Hubbard, 2001). The more important evidence for the nativist account has been negative, in that no case where synesthestic percepts could be traced to some environmental stimulus had been produced (but see Galton, 1880). There is also data that suggest that synesthesia may be inherited (Baron-Cohen et al, 1996), leading some researchers to argue that this rules out any learning component (Ramachandran & Hubbard, 2001). This logic is unconvincing since any predisposition towards synesthesia does not necessarily answer the question of how particular inducer-concurrent pairings arise.

In the first section I present a detailed case study of a grapheme-color synesthete, AED, who has derived her letter-color associations from a childhood toy. Among the results are data showing that AED's synesthesia seems to arise relatively early in visual processing as the perceived brightness of her colors as measured by color matching is altered by embedding letters in lightness illusions (Adelson, 2000). Furthermore, her synesthestic colors can be changed in a systematic way by manipulating the similarity of letters (by changing the font) to those found in the original toy suggesting specificity in her representation of the inducing stimuli.

Color terms and color perception

Does linguistic categorization change the visual system or are differences in behavior a result of the online use of language in tasks?

The evidence from synesthesia suggests that long lasting and profound changes in perception can result from learning. Synesthetes, however, represent a special population and while there is support for the view that they represent an extreme of ordinary function, the underlying mechanisms remain unclear (Maurer, 1997; Ward et al, 2006). Moreover, the label synesthesia may (and probably does) group together disparate phenomena (Ward & Mattingley, 2006). One might then ask whether or not experience plays a role in perception in more ordinary cases?

One much debated possibility is whether or not linguistic categorization influences color perception (Berlin & Kay, 1999, Kay & Kempton, 1984. This topic is of wide interest in part because much about basic color perception (i.e. matching adjacent color patches using lights) is well understood and relatable to the physiology of the retina (Wandell, 1995). Despite the relative universality of cones and early retinal processing, languages in the world have been found to vary with respect to color naming (Davidoff et al, 1999). While some researchers have proposed that this variation can be reduced to underlying and universal regularities (Bornstein et al, 1976; Regier et al, 2005), this would still leave open the question of whether or not the demands of communicating about the world with other people alters color perception, with colors having the same name becoming more similar or less discriminable or colors having different names more discriminable, or both (Ozgen & Davies, 2002). One way of interpreting these ideas is to claim that the habitual need to partition a continuous region of color space as say blue or green has led to alterations in perceptual processing akin to the categorical perception found in phoneme perception. A finding that color perception is altered by linguistic categorization would provide evidence for a strong version of the Whorfian hypothesis in which perception is determined by language.

An alternative to the perceptual learning hypothesis is the possibility that language may play a role online in making simple perceptual judgments (Kay & Kempton, 1984; Roberson & Davidoff, 2000). Some prior work has supported the online account, but has relied on paradigms with obvious memory components and or subjective judgments such as deciding which of three colors is most dissimilar to the other two (Kay & Kempton, 1984). Section 2 presents a series of experiments that advance the online hypothesis by showing that some color judgments where stimuli are simultaneously available and where the judgment is objective (i.e. there is a right and wrong answer) are selectively affected by a linguistic dual task. These results do not obtain in all cases however, and arranging the stimuli so that there is no spatial comparison necessary eliminates the effect.

Face adaptation and face spaces

How does our judgment of the same stimulus (say a face) change over time, and can we use those changes to say something about visual representations and processes?

The first two lines of research above tested relatively permanent effects of experience on perception. However, profound but short lived changes in perception can also be induced relatively quickly. In adaptation paradigms, prolonged exposure to a stimulus such as motion or oriented lines, produces changes in the perception of subsequently presented test stimuli, with the general result that judgments of tests similar to the adaptor are biased away from the adapting stimulus. The direction and magnitude of these changes can be measured and used as a way to explore the underlying representations, for example, by showing that stimuli are processed in channels vs. spaces, or whether or not there are neurons tuned to individual features, conjunctions of features, or both (Suzuki, 2005; Held, 1980). Adaptation is widely used in vision encompassing a wide variety of stimuli ranging from oriented lines to point light walkers (Gibson & Radner, 1937; Troje et al, 2006).

A relatively new line of work has shown that adaptation effects can also be obtained using faces, with aftereffects in the perception of gender, identity, emotion, and ethnicity (Webster & MacLin, 1999; Leopold et al, 2001; Webster et al, 2004). Some researchers following up these findings have argued that identity based adaptation relies on an opponent process system analogous to that of color (Leopold et al, 2001). Faces are located in a continuous face space of unknown dimensionality and identified by the direction of the vector which originates at the center of the space (or the average face). This theory contrasts with exemplar based face spaces in which face identity is coded by the values on the various dimensions and no special status is give to the average face or to trajectories which pass through it (Valentine, 2001; Lewis, 2004). One strong prediction of the prototype referenced proposal is that aftereffects will have a particular direction which goes through the origin of the space. (Rhodes et al, 2006; Tsao & Freiwald, 2006) Experiments using computer generated faces based on 3d-measurements from real faces provided support for this view, showing that judgments of the identity of a neutral stimulus are biased away from the adapting stimulus in the direction predicted by the space (Leopold et al, 2001). Somewhat surprisingly, similar behavioral results have been obtained using monkeys, as well as single unit data which shows increased population activity as faces deviate from the average (Leopold & Bondar, 2005; Leopold et al, .2006).

The set of adaptation experiments presented here use a novel objective recognition task to show that identity aftereffects can be obtained with natural face stimuli and no training. Significant increases in recognition performance are obtained using both unknown and famous faces as adapting stimuli and seem to transfer somewhat across viewpoints (side view to front), while adapting to inverted faces produces no effect. While visible identity based aftereffects are readily observed in test faces, it is not certain from people's subjective judgments whether or not the other identity is really more visible, or the better of two alternatives. The findings here show that the aftereffects are specific enough to improve recognition performance. The results also demonstrate that aftereffects can be obtained along what would be multiple paths in a putative face space simultaneously. While these findings do not necessarily contradict the notion of a face space, they do suggest that the face which lies at the origin of the space is not special leaving open the possibility that identity adaptation effects may be explained by exemplar models as well.

Chapter 2: Learning and Synesthesia

The following chapter originally appeared as:

Witthoft, N. & Winawer J (2006). ***Synesthetic Colors Determined by Having Colored Refrigerator Magnets in Childhood. Cortex***

Introduction:

The origin of the relation between what are sometimes termed the inducer and concurrent (Grossenbacher and Lovelace 2001), meaning the stimulus that produces synesthesia and the synesthesia itself, has provoked much debate. One general line of argument is that the relation between inducer and concurrent results from pre-existing mappings between sensory areas that are overactive or fail to be pruned during development (Ramachandran and Hubbard, 2001a). To bolster this view, researchers have pointed to what seem to be innate alignments between sensory dimensions. For example, normal subjects will rate higher pitched sounds as brighter than lower pitched sounds, a trend which seems to hold true for pitch-color synesthetes (Marks, 1975). It has been also been suggested that infants may be innately synesthetic, with sensory differentiation coming only with development and the gradual pruning of connections (or at least development of inhibition) between sensory areas (Maurer, 1993).

Another theory is that learning can influence the development of particular inducer-concurrent pairings. This idea has a long history, beginning over 100 years ago with Galton's pioneering account of a synesthete who perceived dates as colored and claimed that the colors were derived from a history book used for childhood instruction (Galton, 1880; Galton, 1907). Despite the intuitive plausibility of a learning account, this view had largely fallen out of favor due to the inability of researchers to produce even a single case where learning could be traced to a particular environmental source (Marks, 1975; Ramachandran and Hubbard 2003b).

However, recent findings suggest a need to acknowledge a role for environmental influences. Mills and colleagues report a color-grapheme synesthete, MLS, who perceives colors in response to both English and Cyrillic letters (Mills et al., 2002). While the origin of MLS's letter-color mappings are unknown, MLS surprisingly developed new concurrents for English letters following brief laboratory exposures to letters displayed in colors that did not match any of the colors normally elicited by English letters. These new and inadvertently acquired photisms matched those colors given in the lab but were weak and transient, disappearing within a few weeks. A different kind of environmental influence is at work in a recently reported case of a word-taste synesthete, for whom a number of factors, including semantics and phonology, seem to have influenced the link between inducer and concurrent (Ward and Simner, 2003). For example, words containing the phoneme "k" tended to elicit the taste of foods with the same phoneme, such as "cake". The authors argue that the mappings between inducer and concurrent can be conceptually mediated rather than simply reflecting hardwired connections between sensory modules.

This paper presents a series of results from an investigation of a color grapheme synesthete, AED. She has had synesthesia for as long as she can remember and reports all achromatic text as having colors overlaid on the surfaces of the letters or numbers. She is highly educated and has received a PhD in a scientific field. She has no history of neurological or psychiatric disorder and has normal color vision as assessed using "16 Plate Pseudoisochroamatic Color Vision Test" from Goodlite. None of the other members of her immediate family reports having synesthesia of any kind. What sets AED apart from previously reported synesthetes of this type is evidence suggesting that the inducer-concurrent relationship was learned from a refrigerator magnet set and subsequently transferred to the Cyrillic alphabet. Further experiments shed light on the representations of the inducer and concurrent and where in the respective streams of processing they might arise.

Consistency over time

Like other synesthetes, AED reports that her particular set of inducer-concurrent pairings has always been the same. Consistency over time is ordinarily considered a prerequisite for establishing the genuineness of synesthesia (Baron-Cohen et al., 1987; Grossenbacher and Lovelace, 2001; Rich and Mattingley, 2004). We quantified the consistency of AED's photisms by correlating the hue, saturation, and brightness of her matches generated on a computer during two tests 3 weeks apart.

Methods

Stimuli were presented using a Macintosh G4 computer with a 17-inch LCD flat panel monitor using Vision Shell stimulus generating software. Testing was done on two separate occasions, 21 days apart, in a quiet darkened room with the same computer and monitor used for both sessions. The graphemes appeared in gray and AED adjusted the hue, brightness, and saturation of the grapheme using controls also on the screen. Moving the pointer on the surface of a color wheel or brightness slider simultaneously adjusted the color of the letter or number. There were 360 possible hues, 100 levels of saturation, and 128 levels of brightness, encompassing the full range afforded by the monitor. AED had unlimited time to make the matches but could not go back once a match was made. During each matching session AED was presented in random order with the digits 0 to 9 and all letters of the alphabet both uppercase and lowercase (n=62). The consistency of AED's synesthetic pairings was assessed by separately correlating the hue, brightness, and saturation from one test with the same measures from the next. Since the data were not normally distributed on some dimensions, the non-parametric correlation coefficient Spearman's rho is used to assess the strength of the correlations,

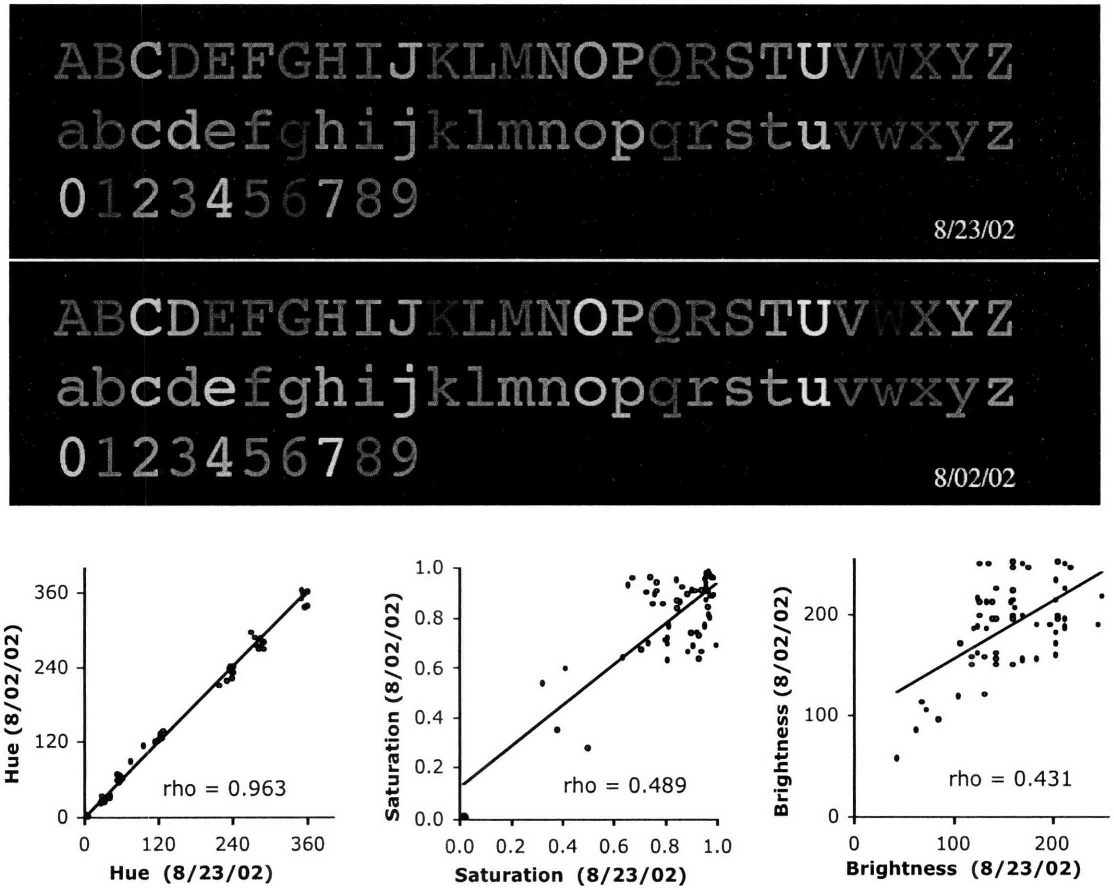


Figure 1. Consistency over time. Top: AED's synesthetic colors as produced in a matching task on two different test sessions. Bottom: Correlations between the two test dates on the dimensions of hue, saturation, and brightness. While all three are highly significant, the hue shows almost no variation between the two dates. Lines represent least square regression.

Results:

The hues of the individual letters taken from the two tests were highly correlated (rho = 0.963, $p<0.001$), while the saturation and brightness were more variable (saturation: rho = 0.489, $p<.001$; brightness: rho =.431 $p < 0.001$, Figure 1). AED clearly satisfies the criteria for this measure of synesthesia. While all of these correlations are highly significant, it is notable that

AED is most specific about the hue dimension than either saturation or brightness. AED did this task very quickly, doing all the trials in 15 minutes or less on each test.

Refrigerator Magnets

Reorganizing AED's alphabet (Figure 2) reveals that with the exception of the B, every sixth letter has a similar hue. When asked about the pattern AED recalled having a toy as a child with similarly colored letters. A photograph of the toy, obtained from her parents, appears in Figure 3. The relationship between AED's synesthetic colors and the refrigerator magnets is obvious and we conclude that AED learned her colors from this toy. As far as we have been able to determine, this represents the first and only documented case of experience determined color-grapheme synesthesia. It also presumably extends the consistency of her inducer-concurrent mappings across nearly 30 years.

An outstanding question remains regarding the origin of AED's synesthesia for numbers. AED believes that the numbers were also learned from an accompanying set of refrigerator magnets that included numbers and basic mathematical operators ('+', '-', '/', '=') for which AED also experiences synesthesia. However, AED is no longer in possession of the set and we have been unable to locate one and are therefore unable to say anything definitive about the origins of these colors.

Figure 2. Patterns in AED's synesthetic alphabet. Aligning AED's matches from experiment 1 so that every 6th letter falls in the same column reveals a striking pattern. With the exception of the letter 'B' and possibly 'I', every 6th letter has a similar hue.

Figure 3. Experience-dependent synesthesia. Top: AED's synesthetic colors, as adjusted by her in random order using a color wheel and a brightness bar. Bottom: AED's childhood refrigerator magnet set, recovered from her parent's attic. Note that the "B" was apparently missing in her childhood and recently found, and the "G" and "H" were present during childhood and recently lost.

Transfer to Cyrillic

AED's experience dependent mapping of inducer to synesthetic color extends beyond the English alphabet. AED moved to Russia at age 3 and developed concurrents for Cyrillic letters. Many of the letters in the Cyrillic alphabet visually resemble those used in English, even while standing for different sounds or concepts. For example, the symbol '3' in English is a number, but in Russian, a very similar symbol is '3' is read 'z' as in zoo. However, there are also letters in the Cyrillic alphabet that are not visually similar to any letters in the English alphabet but represent phonemes that also occur in English. We investigated whether or not there were any correspondences between the colors elicited by the two alphabets and whether visual or phonetic or even conceptual similarity might prove to be a determining factor.

Methods:

Colors for Cyrillic letters were obtained using exactly the same environment, software, computer and monitor as the matches for English letters. A total of 62 stimuli (31 upper case and 31 lower case) were presented. Twenty of the 31 pairs of Cyrillic letters have the same basic form as English letters or are mirror reversals of English letters and were considered visually similar to English letters. There was no ambiguity about whether the letters were English or Cyrillic both because AED knew that all letters on this test date were Cyrillic, and because there are subtle differences between even the English and Cyrillic letters classified as visually similar. Of the remaining 22 letters, 10 (5 upper case letters and their 5 lower case counterparts) had the same sound as single English letters. The last 12 did not clearly resemble any particular English letter and represented sounds that are not used in English or that required more than one English letter to represent (for example, sh) and are not included in the analysis.

Visual Similarity			Phonetic Similarity		
И	ee in see	N	Ф	ef in face	F
Я	ya in yard	R	П	p in pot	P
З	z in zoo	3	Г	g in go	G
А	a in car	A	Л	l in lamp	L
Р	trilled r	P	Д	d in do	D
Х	ch in loch	X			
б	b in bit	b			
М	m in my	M			

Figure 4. Transfer of synesthesia to Cyrillic. In the right column are some examples of Cyrillic letters that have the same basic form or are mirror-reversed versions of English letters. There were 40 such examples total (20 upper and 20 lower case). There were 10 letters with phonetic but not visual analogs to English letters (5 upper and 5 lower case), shown on the right.

Results:

Generally, the concurrents for Cyrillic letters depend on their visual similarity to English letters (Figure 4). Correlating the hue, value and saturation of the 40 visually similar letters with the average of their counterparts generated in the two matching experiments with the English alphabet gave significant correlations for hue and saturation but not brightness (hue, rho = 0.960 ; saturation, rho = 0.611, both $p<0.001$). The correlation among the hues is nearly perfect despite the fact that many of the Cyrillic letters stand for different sounds or concepts. This suggests that visual similarity to English letters (or the original letter set) is a dominant factor in determining the synesthetic color generated by Cyrillic and has implications for the representation which induces AED's synesthesia.

Comparing the Cyrillic letters that have only a phonetic similarity to English letters also revealed a near perfect correlation with respect to hue (rho = 0.973, p<0.001) but no significant correlations for either saturation or brightness (Figure 5). These results are nicely consonant with a previous report of an English speaking synesthete, MLS, who developed photisms for Cyrillic letters after learning Russian in high school (Mills et al., 2002). Like AED, MLS also showed similar colors for most visually similar letters in the two languages, and colors based on phonetic similarities for the majority of the rest.

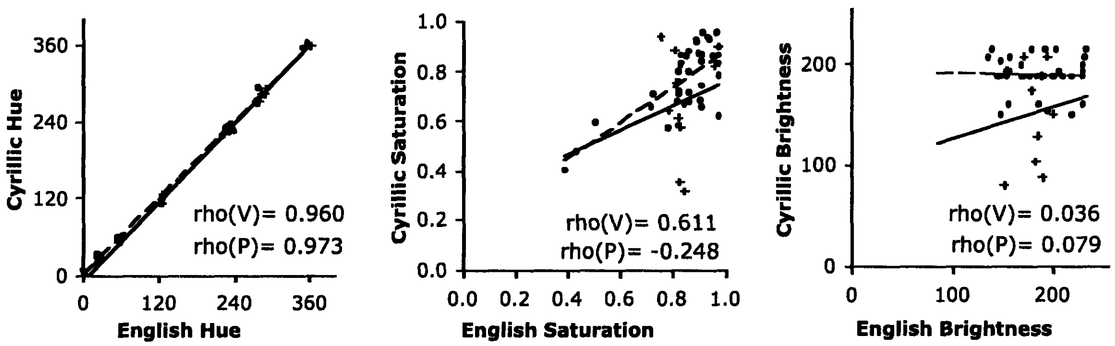


Figure 5. Synesthesia for Cyrillic letters. Pictured above are the correlations for hue, saturation, and brightness between the Cyrillic letters and the average values on these same dimensions taken from English letters. Visually similar letters (dashed lines) and phonetically similar (solid lines) are least square regression lines. The hue for the Cyrillic letters correlates perfectly with those from English even when the relationship is only phonetic. The saturation shows significant correlation only when the Cyrillic letters are visually similar to English letters.

What is the level of representation of the inducer?

At first it might seem strange that phonetic similarity in the absence of visual similarity can have an effect. But the relationship between Cyrillic letters that have no visual counterpart in English is no different than the relationship between many lowercase letters in English and their uppercase partners. For example, the relation between 'a' and 'A' is analogous to the relation between 'Ф' and 'F', in that the only relation they have is they stand for the same sound.

It has been reported that for some color-grapheme synesthetes the case of the letter or even the font may have some effect on synesthesia, with more prototypical fonts yielding more vivid synesthesia (Ramachandran and Hubbard, 2003a; Ramachandran and Hubbard, 2003b). Others have suggested that for most synesthetes font and case have no impact on the color (Grossenbacher and Lovelace, 2001). This question is of interest because noting what type of changes in the stimulus produce changes in the concurrent provides insight into the specificity of the representation of the inducer. For example, if synesthesia were only produced by one particular font, then it could be argued that the inducer is a representation that contains very specific shape information. If variations in font or case have no effect on the concurrent than it is likely that the representation that produces the concurrent is more abstract, concerned with the category to which the letter belongs. The method of obtaining synesthetic colors used here is useful to this end. Previously, most researchers have used color names to verify consistency over time, but this approach may mask significant differences in the concurrent since changes in case may result in subtle but detectable changes in the concurrent that are lost when mapped onto the coarser code of language.

Since hue was clearly unaffected by changes in case or shape, we investigated whether case or font had any systematic effect on the saturation or brightness of AED's synesthesia. In the English alphabet some lower-case letters are visually similar to upper-case letters, distinguished from them largely by their relative size or position ('W' vs. 'w') while other letters can have the case determined by shape alone ('A' vs. 'a'). To see if visual identity was an important factor with English letters, we also looked at the importance of visual similarity of upper to lowercase letters on both saturation and brightness. For the Cyrillic alphabet, we separated the letters both by case and by similarity to the English capital letters and examined the impact on saturation and brightness.

Methods:

Brightness and saturation data from the English letters were taken from the consistency experiment (in Times font) and from an additional testing session using the Sand font (a cursive font). The data were analyzed in a 2x2x2 ANOVA, with case (upper vs. lower) and font (Times vs. Sand) as repeated measures over 25 letters and similarity between the cases as a between subjects factor. The letter "I" was excluded because it was mistaken for the number "1" in one session. The matching results for the Times font were averaged from the two test sessions, while the Sand results came from a single session. Eleven letter pairs were considered visually similar between cases and 15 pairs were considered dissimilar.

Responses from the Cyrillic matches were analyzed in a 2x2 ANOVA with case as a within subject factor and similarity to the English capitals (visual vs. phonetic), as a between subject factor. Several of the capital Cyrillic letters are visually similar to lowercase English letters resembling either 'b' or and upside down 'h'. Since it was unclear how to treat these letters they were excluded from the analysis. This left 17 pairs of letters that were visually similar to English, and 5 pairs that were phonetically similar.

Results:

No significant effects for brightness were found in any contrast using either the Cyrillic or English letters. Saturation of the letters showed a number of effects. For the English letters there was a main effect of case with AED rating the uppercase letters as significantly more saturated than the lowercase (mean saturation of uppercase = 0.824, SE = .019, mean lowercase 0.775, SE = .020, F(1,23) = 11.983, $p<0.01$). There was also a main effect of font, with Times more saturated than Sand (mean Times = 0.837, SE = .018, mean Sand = 0.762, SE = .020, $F(1,23) = 18.580$, $p<0.001$). These factors also interacted, such that the effect of case was bigger for Times than for Sand ($F(1,23) = 7.131$, $p < 0.05$), with all the saturation values in Sand resembling the lowercase values in Times (Figure 6). There was no main effect of similarity, but there was an interaction between similarity and case, such that the uppercase was more saturated than lowercase for pairs that were visually dissimilar, but not for pairs that were similar ($F(1,23) = 6.123$, $p < 0.05$). There were no other significant interactions for the English letters.

For the matches generated to Cyrillic letters there was also a main effect of case on saturation ($F(1,20) = 6.211$, $p = 0.022$) with the uppercase letters showing significantly greater saturation than their lowercase counterparts. There was also an interaction between the case and similarity of the letters to English ($F(1,20)=5.94$, $p=.024$). This was entirely due to a reduction of saturation for the lowercase Cyrillic letters visually similar to English. The remaining Cyrillic letters, those that are not visually similar to English letters, showed no effect of case on perceived saturation and in fact overall had a similar saturation to the lowercase English letters (Figure 6, right panel). The fact that both upper and lowercase letters in these Cyrillic letters were reduced in saturation relative to their English counterparts further supports the relevance of visual information to the inducer representation.

Overall, the pattern seems to be that greater visual similarity either to the original refrigerator magnet set or possibly to a prototypical font representation causes AED's synesthesia to appear more highly saturated. Thus the more standard font (Times) led to more saturation than the non-standard font (Sand), and uppercase more saturation than lowercase. Moreover, the loss of saturation for lowercase letters was greater for those that are not scale or position transformations of their uppercase counterparts.

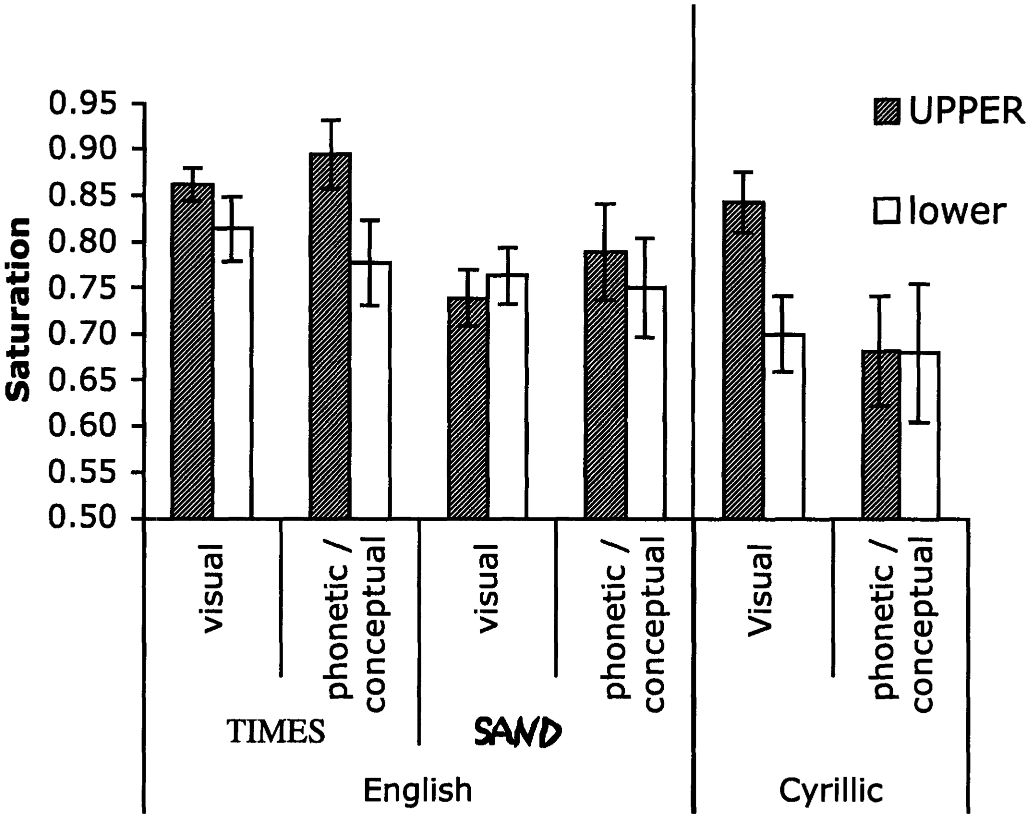


Figure 6. Saturation for AED' s synesthesia. Left and Center panels: English letters. Upper case letters are more saturated than lower case letters in the standard **TIMES** font (left), but not in the more unusual **SAND** font (center). Generally, the letters in Sand have a saturation similar to that of the lower case Times letters. Moreover, the reduction in saturation with lower case compared to upper case is more evident in those letter pairs for which the upper and lower case are not visually similar to each other. These findings all support the notion that letters which are less similar to a prototype are also less saturated. The letters C, K, O, P, S, U, V, W, X, Y, Z, were designated as being visually similar in upper and lower case, whereas the letters A, B, D, E, F, G, H, J, L, M, N, Q, R, T were designated as being visually dissimilar between upper and lower case. Right panel: Effects of similarity to English capital on the saturation of AED's concurrents for Cyrillic letters. Visually similar upper case letters are highly saturated, like the upper case Times letters in English. Their lower case counterparts, as well as letters that are only phonetically or conceptually similar to English letters (whether upper case or lower case), all show a reduction in saturation. As with English letters, the more the letters differ from an upper case English prototype, the less saturated they appear.

One interesting and compelling possibility is that visually dissimilar letters activate synesthesia only indirectly. For example, looking at a lowercase letter may activate neural regions that sufficiently overlap or spread activation to those that would be activated by the uppercase letter. This overlap in pattern of activation is similar enough to evoke synesthesia, but the difference results in weaker saturation.

Letters in Cyrillic standing for sounds that also occur in English would also weakly activate their counterparts either through spreading activation or neural overlap and subsequently generate the synesthetic concurrent. On the other hand, it could be that multiple levels of representation of the inducer are each capable of activating the concurrent directly on their own. How they do so differs as the representation changes from visual to phonetic or conceptual.

What is the level of representation of the concurrent?

The specificity with which AED selects hue, and the subtle changes in saturation with changes in case or visual similarity suggest that AED's concurrents arise early in the stream of visual processing. Certainly the hue seems to arise at a level earlier than linguistic categorization. Nonetheless, skeptics might argue that AED may just have an extraordinary memory for particular color-letter pairings as could be learned by any person willing to put in the time.

Not all color-grapheme synesthetes claim to see colors as part of the visual scene; some report seeing the colors only in their mind's eye and some others say only that they have very specific ideas about what colors letters are (Grossenbacher and Lovelace, 2001). Synesthetes like AED, who report seeing colors on the surfaces of letters (sometimes referred to as projectors; Smilek et al, 2001) are of particular interest because the ways in which their concurrents interact or fail to interact with normal visual processing can point to where in the stream the percepts arise (Ramachandran and Hubbard 2001; Palmeri et al., 2002). While AED's subjective experience is ultimately impenetrable, and therefore we cannot talk conclusively about what

AED ‘sees’, it is possible to infer whether the concurrent is processed by very early visual mechanisms, mechanisms that would ordinarily be considered insulated from information in associative memory.

To this end we tested whether AED’s synesthesia was affected by lightness constancy illusions. Working with lightness constancy can reveal whether synesthetic colors are integrated into early stages of perceptual processing and treated as part of the visual scene. In the Checkershadow illusion (Adelson, web.mit.edu/persci/people/adelson/checkershadow_illusion.html), two identical squares marked A and B appear to be very different shades of grey (Figure 7, left). The illusion is created because the visual system takes into account illumination conditions when computing the reflectance properties of a surface. Identical amounts of light are reflected by the two squares, but for something apparently in shadow to reflect as much light as something that is directly lit, the surface in shadow must have greater reflectance than the lit surface. The visual system thus ‘discounts the illuminant’ and interprets surface B as brighter than surface A (Adelson, 2000). The same process is at work in the Snakes illusion (Adelson, 2000) except that the illusion results from B appearing to lie beneath a dark transparent surface instead of beneath a shadow (Figure 7, right).

We reasoned that if AED’s photisms were incorporated into the normal stream of visual processing at an earlier stage than lightness constancy is computed, synesthetic colors produced by achromatic letters embedded in the illusion should also be adjusted by the same constancy mechanisms. That is, given that the visual system detects two colors of identical luminance, but one seeming to be in a shadow, lightness constancy mechanisms should determine that the one in the shadow must actually be brighter. It does not matter that those colors are not in the physical stimulus, only that the brain treats them as real, incorporates them into its representation of the visual scene and processes them. In short, if AED generates matches for letters presented in a shadow or beneath a transparent surface that are brighter than the same letters when perceived to be directly lit, her concurrents arise before lightness constancy has been fully computed.

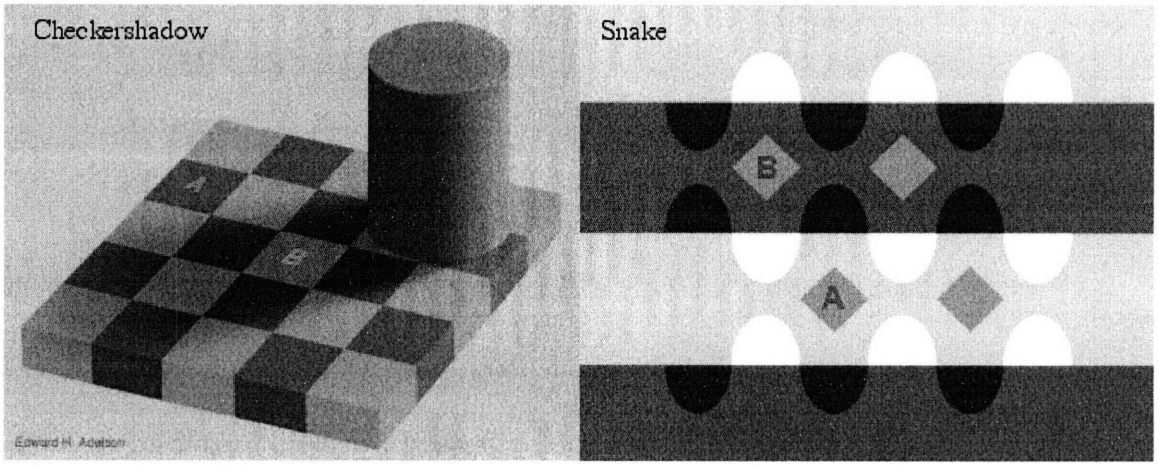


Figure 7. Lightness constancy illusions. In the Checkershadow illusion (left) and the Snakes illusion (right) the regions marked A and B are of identical lightness. However, lightness constancy mechanisms "correct" the apparent brightness such that the two regions marked B appear brighter either because they are in apparent shadow (left) or because it is behind a dark transparent surface (right). For the lightness experiments, AED was presented with gray letters, one per trial, either in region A or B, either in the Checkershadow or the Snakes illusion. AED adjusted the color and brightness of a small rectangular patch outside the illusion until it matched the apparent color and brightness of the synesthetic color. The two illusions were used on separate test dates. Both illusions are creations of Ted Adelson.

Methods:

The software from the previous color matching tasks was modified so that achromatic letters now appeared embedded in one of the two lightness illusions. On two test dates, AED did the task using letters embedded in the Checkershadow illusion, and on a third occasion did the matching on the Snakes illusion. Each capital letter appeared in two trials, once on the square in the light and once on a square in the shadow or beneath the transparency (marked "A" and "B", respectively, in Figure 7). AED adjusted the color of a patch located above the illusion using a color wheel and brightness slider until the test patch matched the color of her synesthesia. All other conditions were identical to those in the color matching tasks.

To quantify the effect of the lightness illusion on real colors, 3 non-synesthetic control subjects with self-reported normal color vision also did the Checkershadow task, except that the letters appearing in the illusion had colors. The control subjects matched the test patch outside the illusion to the color presented on screen. AED did a similar control task to make sure her lightness constancy was normal, except that instead of letters she was presented with ovals embedded in the illusion. The colors for the ovals (for AED) and the letters (for non-synesthetic control subjects) were generated from AED's previous matching results.

For both AED and controls each experiment using letters had 50 trials with each letter or oval appearing twice (once in each location for 25 pairs). Only capital letters were presented and the letter 'I' was excluded. For the experiment using colored ovals, AED was given 30 trials (15 pairs of colors). Data were analyzed by doing a one-sided paired t-test on the brightness values of each of the letters (or ovals) taken from the two different locations in the illusion. By hypothesis, it was expected that stimuli appearing in the shadow would be rated as brighter than those appearing in the directly lit square. Brightness values are units of pixel intensity from 0 to 255.

Results:

AED showed a highly significant effect of the Checkershadow illusion on the perceived brightness of her photisms ($n=25$, mean difference 17, $SE=3.4$, $p<0.001$), demonstrating that her concurrents are available to the early stages of visual processing. A similar result was found with the Snakes illusion; AED's photisms were systematically brighter when perceived to be beneath a dark transparent surface than directly lit (n=25, mean difference 14, SE=5.4, $p < 0.01$). As expected, normal controls and AED showed similarly large and systematic effects in the predicted direction when matching real colors (AED: n = 15, mean difference 37.3, SE=6.7, $p< 0.001$; controls: n =25, mean difference 40, SE=4.5, $p<0.001$). Thus in all cases, the brightness of the colors, whether real or synesthetic was modulated by the apparent illumination.

Analysis of a control experiment in which the brightness of achromatic letters was systematically varied with no change in the apparent illumination showed no correlation between the brightness of the achromatic letter and AED's matches. This suggests that the effects described above are due to lightness constancy mechanisms operating directly on AED's synesthetic colors, and not due to the apparent difference in brightness of the inducing letter in the different regions of the illusions. That is, the cues to shadow (or transparency) surrounding the letter affect the apparent brightness of the synesthesia, but the brightness of the inducer itself does not matter.

An obvious question is why the effect on AED's synesthetic colors is smaller than for real colors. One possibility is that lightness constancy is the result of several mechanisms and only a subset of these processes has access to the synesthetic colors. We presume based on these experiments that the synesthetic color is added to the visual scene via feedback from letter recognition areas to early stages of visual processing involving color. Since simultaneous contrast has been shown to influence neural activity as early as V1 (MacEvoy and Paradiso, 2001) and given that there is a simultaneous contrast component to the illusions, it is likely that the synesthetic color is not available to some of the earliest constancy mechanisms.

Discussion

Together these findings document the first clear case of experience-dependent color-grapheme synesthesia. For AED, exposure to colored refrigerator magnets in childhood led to a lifelong pairing between graphemes and synesthetic colors. However, we do not claim that the refrigerator magnets caused AED's synesthesia, only that they shaped its manifestation. There are many reasons to think that synesthesia does not require refrigerator magnets or the like. For example, studies of synesthetes and their relatives have suggested that synesthesia is heritable with women much more likely to be synesthetic than men (Baron-Cohen et al, 1998). Thus while it may not be the case that all synesthetes, or even the subset with the same type of synesthesia as AED have learned their particular inducer-concurrent pairings, it may be that for those who will become synesthetes, there exists a sensitivity to certain kinds of input (in this case colored letters) that can result in this kind of learning. This adds to a growing body of evidence that experience can play a role in determining inducer-concurrent pairings (Ward and Simner, 2003; Mills et al., 2002).

The fact that AED's synesthesia was learned from a magnet set does not mean that one can dismiss it as an ordinary high level memory association. In fact, the lightness constancy experiments make a strong case that AED's synesthesia is available to early visual processes which are considered encapsulated with respect to associative memory. In short, this study highlights the fact that a learning account of synesthesia can be consistent with an early processing account. These results are quite surprising because they imply that synesthetic color is not added to the end-product of visual form recognition. Rather the color seems to be added to the processing stream such that it becomes part of the stimulus subject to further visual processing. Whether this finding generalizes to other synesthetes is of course an important question. Our results suggest that careful case studies can point to surprisingly tight connections among synesthetes. For example, the fact that AED's synesthesia generalized to Cyrillic through visual similarity where possible and phonetic similarity secondarily is in perfect agreement with the case reported by Mills and colleagues (2002), and the reduction in saturation with lowercase letters and non-prototypical fonts is in agreement with the claims of Ramachandran and Hubbard (2003b).

One way to interpret AED's inducer-concurrent relation is using the model for color-grapheme synesthesia put forward by Rich and Mattingley (2002). It proposes based on work in neuroscience, separate processing hierarchies for form and color. Synesthesia happens because functional connections are formed between different levels of representation in the two pathways. The lightness constancy results, the specificity of the hue, the subtle effects on saturation, and AED's report that the colors appear on surfaces in the world, all suggest that functional connections are made from the form pathway to somewhere early in the color processing stream, prior to abstract categorical representations. The level and content of the representation of the inducer is less clear, though we have enumerated several possibilities in our discussion of the results. However, the usefulness of precise color matching should be clear. Future studies using these methods may provide a way of generating a useful taxonomy of color grapheme synesthetes, based on information about the inducer and concurrent gleaned from careful analysis of the matching data.

Chapter 3: Language and Color Perception

Introduction

The first section of this dissertation looked at the role of associative learning on color perception in synesthesia and showed that at least in some people, learning can alter perception in profound and long lasting ways. Nonetheless, synesthetes are obviously an unusual subset of people and it is uncertain whether the mechanisms of synesthesia are those found in all people, but taken to an extreme, or something entirely separate (Rich & Mattingley, 2002). It is therefore of some interest then to look for more ordinary situations in which categorization or association might affect perceptual processes. One such area which has generated much debate and numerous experiments is the question of whether and how language and linguistic categories affect color perception.

These questions are driven by the observation that language divides the continuous spectrum of color into discrete categories (e.g., in English: yellow, green, blue, etc), and different languages divide up the color spectrum in different ways (Davidoff, Davies, & Roberson, 1999; Kay & Kempton, 1984; Kay & Regier, 2003; Roberson, Davidoff, Davies, & Shapiro, 2004, 2005; Roberson, Davies, & Davidoff, 2000). Do these linguistic categories play a role in how people perceive colors? The issue has received much empirical attention, with some studies finding a role for language in people's color judgments (Kay & Kempton, 1984; Pilling & Davies, 2004; Roberson & Davidoff, 2000; Roberson et al., 2004, 2005; Winawer et al, 2007), while others have found no evidence for the involvement of language (Bornstein, Kessen, & Weiskopf, 1976; Franklin, Clifford, Williamson, & Davies, 2005; Franklin, Pilling, & Davies, 2005; Regier, Kay, & Cook, 2005). Note that the question of whether and how language might affect color judgments is somewhat distinct from the question of how color categories arise and whether or not they follow some universal pattern.

In this paper we ask whether language can play a role even in simple, objective, low-level perceptual decisions about color. For example, do linguistic processes play a role even in basic decisions such as judging whether two squares of color are exactly the same? Further, if it is the case that language shapes the way its speakers see colors, how does it do it?

There are at least two separate putative sources for the influence of language on people's perceptual judgments. One possibility is that learning and using a language shapes and distorts the underlying perceptual space during development (or during language acquisition more generally) to bring it into closer accord with linguistic categories. For example, colors that fall into the same linguistic category may be moved closer in perceptual space (Goldstone, 1994; Ozgen & Davies, 2002). This type of influence would be an *underlying* or a *long-term* shaping of perceptual space by language.

A second way that linguistic categories might influence perceptual decisions is by language playing an active, *on-line* role during the perceptual decision process itself. This would imply that if linguistic processing is somehow disabled during a perceptual task (e.g., by verbal interference), then the correspondence between linguistic categories and perceptual performance should decrease. Disabling linguistic processing should have no effect on any *underlying* changes in a putative perceptual space, since the proposal is that repeated categorization has changed the output of this process and it is not dependent in any way on language use.

These two possible ways for language to influence perceptual decisions – the *underlying/long-term* shaping and the *on-line/active* involvement - are fully separable. One of these could be true without the other being true. It is possible that linguistic processes help shape the underlying representations in the process of language acquisition, but are not actively involved during the perceptual decisions themselves. Or it is possible that linguistic processes are only involved on-line during perceptual decisions, but the underlying perceptual space remains unchanged by language. It is also of course possible that both of these are true, or neither of them. This paper aims to examine the second possibility – whether or not linguistic processes are involved *online* during perceptual decisions (perceptual decisions that should not, in principle, require access to language).

1.1 Previous Evidence: Similarity Judgments

A number of previous studies have argued for this second possibility, namely that linguistic processes are involved online, and can cause categorical distortions in people's color judgments (e.g., Kay & Kempton, 1984; Roberson et al., 2000). For example, Kay & Kempton (1984) (henceforth K&K) tested English and Tarahumara speakers on colors spanning the blue/green border in English. Unlike English, Tarahumara does not make a linguistic distinction between blue and green. K&K showed five English speakers and four Tarahumara speakers triads of color chips drawn from the blue/green region in color space. For example, subjects might have been shown three chips (let's call them A, B, and C) spanning from green (A) to blue (C), with the middle chip (B) falling on the green side of the blue/green boundary in English. In each triad the subjects were asked to pick out the chip that was most different from the other two. K&K found that English speakers were more likely to pick the chip that fell into a different linguistic category – the blue chip (C) – as being the odd one out. Unlike the English speakers, the Tarahumara were equally likely to choose A and C as the odd one out.

K&K reasoned that because the task presented subjects with a difficult and ambiguous choice without an obvious right answer, subjects may have looked for other clues to help them disambiguate the options. As K&K wrote: "We propose that faced with this situation the English-speaking subject reasons unconsciously as follows: "It's hard to decide here which one looks most different. Are there any other kinds of clues I might use? Aha! A and B are both CALLED green while C is CALLED blue. That solves my problem; I'll pick C as most different." Using this naming strategy would produce the bias observed in English-speakers' responses. Because such a naming strategy would not be available to Tarahumara speakers (because no obligatory linguistic distinction is made in Tarahumara between green and blue), the Tarahumara could not use the same strategy.

To test this proposal, K&K conducted a second study designed to disrupt the English speakers' naming strategy. In the second study, K&K showed the same triads to English speakers, but this time instead of being able to view all three colors simultaneously, subjects were able to see only two colors at a time. The color chips were placed in an apparatus that allowed subjects to see chips A and B or chips B and C (by sliding a window), but never all three chips at the same time. Further, the chips were described by the experimenter in a way that was meant to disrupt the subjects' natural naming strategy. For example, subjects would be shown chips A and B and told "You can see that this chip (experimenter points to A) is greener than this chip (points to B)."

The experimenter would then slide a window on the apparatus to allow the subject to view chips B and C and would say "You can see that this chip (points to C) is bluer than this chip (points to B)." The subjects were allowed to slide the window back and forth to examine the colors (A and B together or B and C together), and were asked "which is bigger: the difference in greenness between the two chips on the left or the difference in blueness between the two chips on the right?" Since in this presentation the middle chip (B) was effectively being called both blue and green, K&K reasoned that a naming strategy would no longer be useful in disambiguating the task. And indeed, they found that in this task English speakers showed a different pattern of results than what they had observed in the first study.

K&K interpreted their data as supporting a mild Whorfian view as follows: English speakers naturally used an unconscious naming strategy, which distorted their color judgments, but when the naming strategy was disrupted, they resorted to purely perceptual judgments. K&K's seminal paper provided an important early test of the influence of language on color cognition and has inspired much follow-up work, interest and debate. The debate has brought out several alternative interpretations of K&K's data. Some researchers have argued that K&K's results are consistent with a completely non-Whorfian view (e.g., Pinker, 1994). Interestingly, it is also possible to argue that these results are consistent with a more Whorfian view than that embraced by K&K. We lay out both of these sets of arguments below.

First, a Whorf skeptic might argue that the differences between English and Tarahumara speakers observed in K&K's first study do not actually demonstrate differences in color perception between the two groups. Some authors have pointed out that the study relied on a very small number of subjects (only 4 Tarahumara and 5 English speakers were tested) (Davies, Sowden, Jerret, Jerret, & Corbett, 1998). Second, the pattern of results from the Tarahumara does not look different from chance performance. It is possible that the Tarahumara simply didn't understand the task, or used a different criterion of exactness when making their selections. A task showing equivalent performance between the Tarahumara and English speakers across a linguistic boundary that exists in both languages would have been useful in dispelling this criticism. Further, even if both of the above criticisms are addressed, there is a separate problem of interpretation. It is possible that the naming strategy English speakers use in this task is not unconscious as K&K suggest, but rather a conscious way of figuring out how to respond in an ambiguous task (Pinker, 1994). English speaking subjects may not see the blue chip C as being more different from the two green chips (A and B), but since they are forced to choose one chip as being more different, and since the difference in name is the only available difference to help disambiguate the task, subjects may indicate C as most different as a way of complying with the experimenter's demands. If this is the case, then the difference between English and Tarahumara speakers would not stem from any perceptual differences, but only from English speakers explicitly falling back on linguistic labels to please the experimenter.

Further, it is possible that the study designed to disrupt English speakers' normal naming strategies produced a different pattern of results from the first study not because naming was disrupted, but because of some of the other differences between the tasks. For example, in the first study subjects freely viewed all three colors together, whereas in the second they were allowed to see only 2 colors at a time, presented in a special apparatus. These differences in viewing and possible other differences in illumination, viewing time, and presentation could have changed the results. A study that controls all these other factors while directly manipulating the subjects' ability to name the colors would help address this criticism.

At the same time, one could argue exactly the opposite from K&K's results – namely that these same results are consistent with a stronger Whorfian view than what K&K propose. Because the study aimed at disrupting English speakers' naming strategies varied from the first study in many aspects of presentation (as described above), one cannot conclude that effects of language were erased in this second experiment. It is possible, for example, that if K&K had tested Tarahumara speakers with all of the same details of presentation, illumination, viewing time and so on as the English speakers in the second study, there may still have been a cross-linguistic difference in performance between the two groups. The data available in K&K leave open the possibility that cross-linguistic differences in color judgments persist even when naming strategies are disrupted.

1.2 Previous Evidence: Memory

Other approaches to the question of cross-linguistic differences in color cognition have focused on the effects of language on color memory (e.g., Lucy & Shweder, 1979; Pilling, Wiggett, Ozgen, & Davies, 2003; Roberson & Davidoff, 2000;). For example, Roberson and Davidoff (2000) tested subjects' ability to remember a color across a brief delay. Subjects were shown a color chip, and following a short delay were presented two new chips (one identical to the one they had just seen and one that was slightly different). Their task was to pick out the correct chip. Roberson & Davidoff (2000) found that subjects were more accurate when the distracter came from a different linguistic color category than when it came from the same color category. For example, after being shown a blue color chip, subjects were more likely to pick out the correct color chip if the distracter chip was green than if it was another blue. Inserting a spatial interference task into the delay did not significantly affect this categorical advantage. A linguistic interference task, however, eliminated the advantage. Similar effects were found when the stimuli were facial expressions instead of color chips.

These results provide a nice demonstration that people naturally use verbal codes in memory even for simple perceptual stimuli in an objective task. However, it is important to note that they do not address whether or not color perception is categorical nor whether or not there have been long term changes in the visual system as a result of perceptual learning. These experiments cannot address the question of categorical perception because there is no sense in which the stimulus differences (between the color chips) have been equated. It does not matter for the design whether or not there is a categorical advantage for between vs. within. The point of the design is that only between category judgments are affected by linguistic interference, that is, it is the interaction of the interference task with the type of judgment that matters. Moreover, selective effects of linguistic interference should not obtain if the only result of linguistic classification is changes in early visual processing (i.e. those that do not use language online).

But these findings also raise a further question. Is it only because subjects had to recollect the first patch of color from memory that linguistic categories played a role in these tasks? Would subjects rely on verbal codes even when the stimuli to be discriminated are all simultaneously available to perception?

1.3 The current design

In the studies presented in this paper we aimed to combine the strengths of the two approaches described above (the simultaneous presentation of colors used by K&K and the objective measure and verbal interference paradigm used by Roberson & Davidoff (2002)) to test whether linguistic information plays an online role in simple and objective perceptual tasks. For example, in Experiments 1 and 2 subjects were shown three color chips simultaneously, arranged in a triad, and were asked to say which of the bottom two color chips was perceptually identical to the chip on top. Subjects performed this task with and without verbal interference. On half of the trials the distracter chip was in the same linguistic category as the two matching chips (e.g., all chips were blue – this would be a within-category trial). On the other half of the trials the distracter chip was from a different linguistic category (e.g., the two identical chips were blue, and the distracter was green – this would be a between-category trial). If people's color judgments show categorical distortions that accord with linguistic categories, then subjects should be faster to respond correctly when the distracter color chip comes from a different linguistic category (the between-category trials) than when the distracter color is in the same linguistic category (the within-category trials). This pattern would correspond to a categorical distortion, or category advantage in color discrimination. Importantly, if such categorical distortions come about because people spontaneously make use of verbal codes in the process of making color judgments, then verbal interference should serve to diminish this category advantage. That is, the categorical advantage (measured as a difference in reaction time for between-category vs. within-category trials) should be bigger in the no interference condition than in the verbal interference condition.

We do find with our stimuli that there is a comparison but must point out that we are making no claims about having equated the stimulus differences beyond choosing equal steps in the Munsell space. Indeed, the existence of a categorical advantage suggests that the Munsell steps are not equal with respect to our task. As in the Roberson and Davidoff (2000) study, the critical factor is the interaction of the type of interference with the type of discrimination, not whether between trials are easier than within. We use categorical advantage merely as a shorthand for summarizing the differences.

This design allows us to use an implicit measure to test for the involvement of language in an objective perceptual discrimination task. The task is objective in that subjects are asked to provide the correct answer to an unambiguous question (which they do with high accuracy). This feature of the design addresses the possibility that subjects only rely on linguistic representations when faced with an ambiguous task that requires a subjective judgment. If linguistic representations are only brought online for the purpose of making subjective judgments in highly ambiguous situations, then effects of language should not show up in an objective unambiguous task with a clear correct answer. Further, the three color patches are all present on the screen simultaneously and remain in full view until the subjects respond. This allows subjects to make their decisions in the presence of the perceptual stimulus and with minimal memory demands. Finally, we rely on the implicit measure of reaction time, a subtle aspect of behavior that subjects do not generally modulate explicitly. While subjects may explicitly decide to bias their decisions in choosing between two options in an ambiguous task, it is unlikely that they explicitly decide to take a little longer in responding on some trials than on others. All put together, this technique allows us to test people's discrimination performance in a simple, objective perceptual task. Further, by asking subjects to perform these perceptual discriminations with and without verbal interference we are able to ask whether any language-consistent distortions in color discrimination depend on the online involvement of language in the course of the task.

In Experiment 1 we find that occupying linguistic resources reduces the categorical advantage (the degree to which subjects are faster on between-category than on the within-category trials). Results of Experiment 2 confirm that this effect is selectively due to verbal interference, as spatial interference does not change the size of the categorical advantage. In Experiment 3, the display is changed so that all three color patches are shown adjacent to each other on the screen such that the color differences can be observed simply by detecting a color edge (i.e. color patches do not have to be compared across space). In this case, no selective effect of linguistic interference is found, providing a case of a perceptual decision that does not seem to be affected by color language.

2. Stimulus Calibration

Roberson and Davidoff (2000) used 7 Munsell color chips of equal saturation and brightness, spanning the blue green border (5B, 2.5B, 10BG, 7.5BG, 5BG, 2.5BG, 10G). Their experiments used both real color chips and computer simulations of these color patches. Since both sets of stimuli yielded the same results, we proceeded with only the computer-simulated Munsell chips. CIE-Yxy coordinates of the stimuli as they appeared on our displays showed little change in the lightness axis (average Y=52), and orderly variation in the hue axes (5B: x=0.217, y=0.284; 2.5B: x=0.221, y=0.298; 10BG: x=0.223, y=0.305; 7.5BG: x=0.229, y=0.326; 5BG: x=0.236, y=0.344; 2.5BG: x=0.243, y=0.367; 10G: x=0.249, y=0.378). Previous work has found that the 7.5BG chip taken from the Munsell color space is the blue-green border (Bornstein & Monroe, 1980; Pilling, Wiggett, Ozgen, & Davies, 2003; Roberson & Davidoff, 2000; Roberson, Davidoff, & Braisby, 1999).

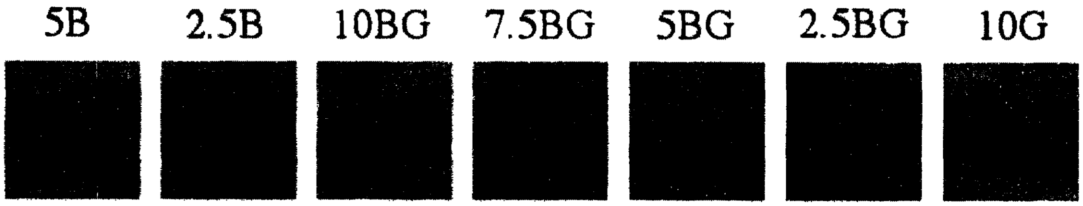


Figure 1: Munsell color chips used to form stimuli in experiments.

We replicated these results in a separate stimulus calibration study designed to identify which among the 7 stimuli is closest to the subjects' blue–green border. 52 English-speaking subjects participated for payment or course credit. Examples of the Munsell chips can be seen in Figure 1. Subjects were shown the 7 chips one at a time, 4 times each in random order. On each trial, subjects were shown a chip and asked to determine as quickly as possible whether the chip was blue or green by pressing the appropriate key. The rationale is that a chip which is on or near the border will be classified more slowly and less consistently than chips falling clearly into a particular category (Bornstein & Monroe, 1980).

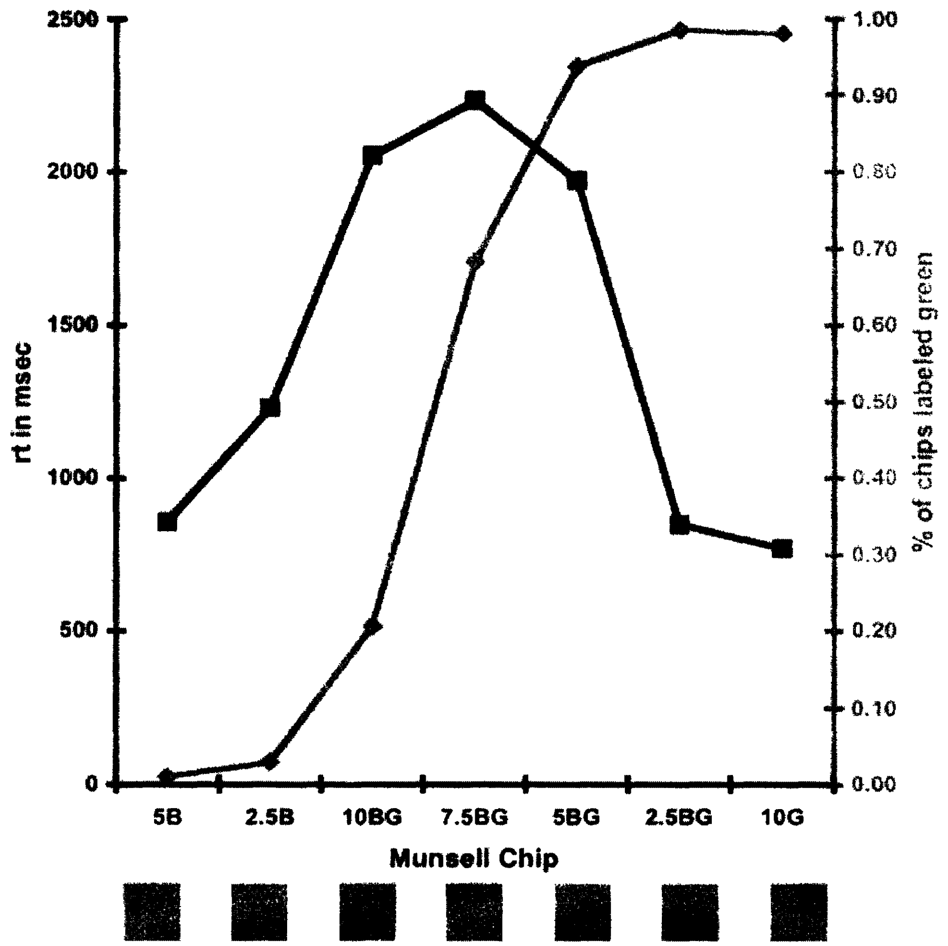


Figure 2. Classification performance of 52 subjects on color patches used in the experiment. The purple line represents the reaction time in ms to determine whether or not a chip was blue or green. The green line represents the percentage of trials the chip was labeled green. Subjects responded most slowly and inconsistently to the 7.5BG chip.

The data (Figure 2) showed that the 7.5BG chip does indeed produce the slowest and least consistent classification, replicating the previous findings, and suggesting again that this is the location of the blue-green border for English speakers. In all of the experiments reported in this paper, comparisons in which both chips were on the same side of the 7.5BG chip were considered within category discriminations. Comparisons of two chips straddling the border or comparisons of any chip to the 7.5BG chip were considered between category discriminations (Figure 3).

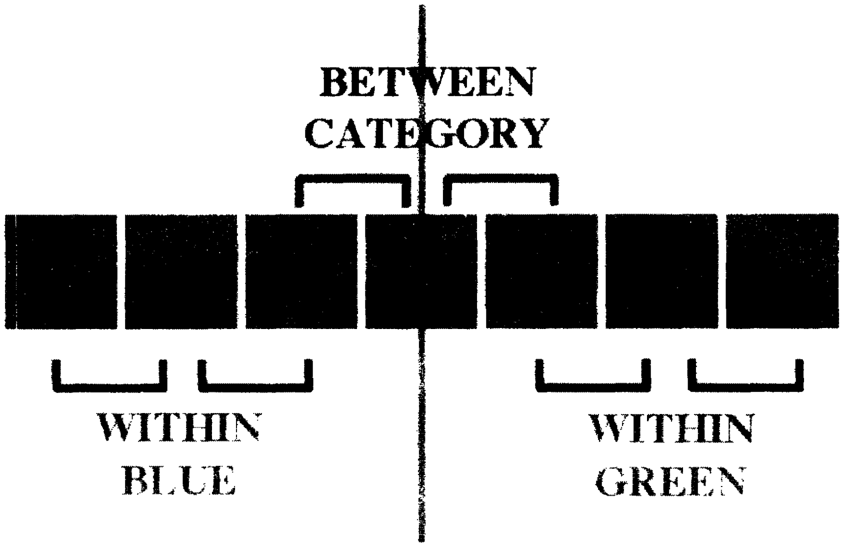


Figure 3. Schematic of discriminations made (for distance 1) in experiments 1, 2, and 3.

3. Experiment 1

3.1 Methods

3.1.1 Subjects

22 native English speakers participated in the study in exchange for payment.

3.1.2 Design & Materials

Subjects completed 192 speeded color matching trials. On each trial, subjects were presented with 3 colored squares arranged in a triad as shown in Figure 4. Subjects indicated which of the two squares on the bottom was identical to the square on the top by pressing one of two keys on the keyboard. If the square on the left was identical to the top square, subjects were to press a designated key on the left side of the keyboard, and if it was the square on the right that was the match, subjects were to press a designated key on the right side of the keyboard. Stimuli remained on the screen until the subject responded. Subjects were instructed to make all responses as quickly and accurately as possible.

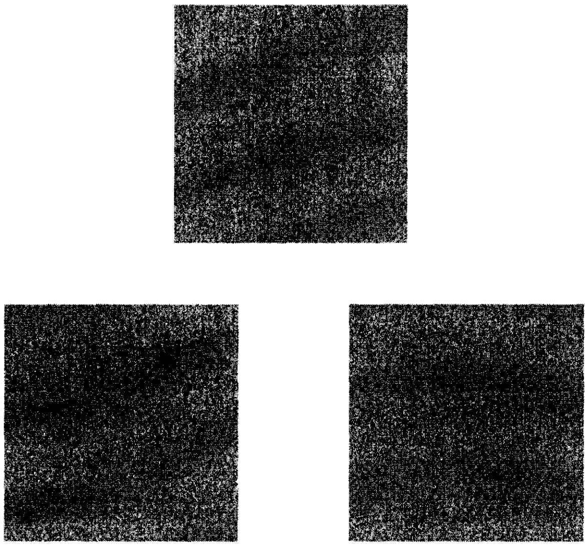


Figure 4. Stimulus display used on color discrimination trials in experiments 1 and 2.

Subjects completed half of the color matching trials under normal viewing conditions, without verbal interference. For the other half of the trials, subjects were asked to perform a secondary verbal task while making their color judgments. All in all, subjects completed 96 trials with verbal interference and 96 trials with no interference. Trials were blocked by type of interference with the order of the blocks randomized across subjects. The no interference block consisted of just the color matching task. For the verbal interference block, subjects were presented with a 7-digit number for 3 seconds and were instructed to rehearse it for a later test. They were then presented with 8 color matching trials identical to those used in the no-interference condition. Following the 8 color matching trials the subjects typed in the 7 digit number. There were a total of 12 digit-strings to recall for each subject within the verbal interference block. Since all of the patches were simultaneously available until the subject responded, reaction time instead of accuracy was used as a measure.

3.1.3 Procedure

Stimuli were presented on an Apple iMac computer on a white background, using the Psyscope 1.2.5 software (Cohen, MacWhinney, Flatt, & Provost, 1993). Subjects were tested in a quiet darkened room.

3.2.1 Analysis

In this and all of the following experiments, only trials on which subjects made the correct color discrimination (and correctly recalled the interference item on trials with interference) were included in the analysis. Following this procedure, an average of 80% of trials for each subject were included in the analyses (SE = 2%). Results were analyzed using a 2 (interference: none, verbal) x 2 (discrimination: between, within) x 2 (color distance: 1, 2 Munsell chips) within-subjects factorial design.

3.2.2 Overview of Results

As shown in Figure 5, subjects showed a categorical advantage when tested without verbal interference: they were faster on between-category trials than on within-category trials. Under conditions of verbal interference, this categorical advantage was eliminated. The details of the statistical analyses are reported below.

3.2.3 Detailed Analyses

There was a main effect of interference, with verbal interference trials producing significantly longer reaction times than no interference trials (mean none = 1045, SE = 79, mean verbal = 1276, SE = 104, F(1,20) = 10.286, p = 0.004). There was also a main effect of color distance, with slower reaction times for the finer discriminations (mean distance 1 = 1360, SE = 113, mean distance 2=960, SE = 63, F(1,20) = 32.783, $p < 0.001$). Finally, there was a main effect of category with between-category discriminations having faster reaction times (M = 1116, SE = 90) than within-category discriminations (M =1205, SE = 85), F(1,20) = 4.432, p =0.047.[1]

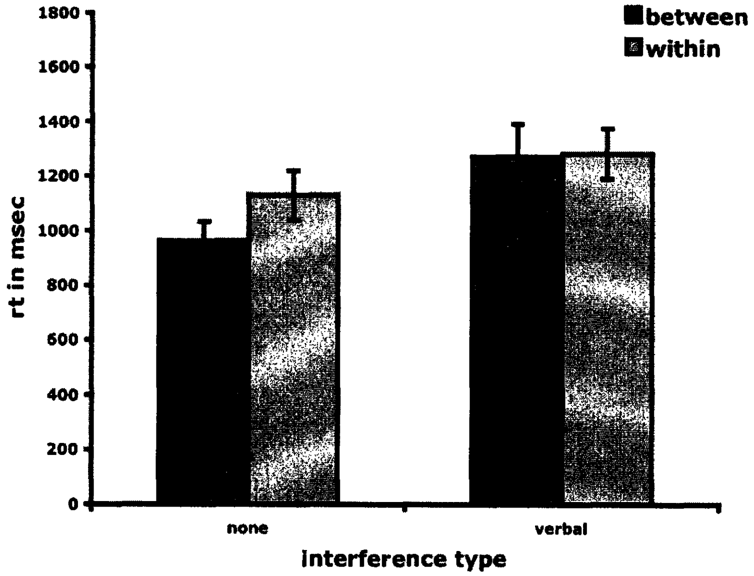


Figure 5. Decline in category advantage from verbal interference in experiment 1. In the no interference condition, subjects are significantly faster on the between trials relative to the within trials. Verbal interference eliminates this advantage.

The effect of category significantly interacted with the effect of distance (F(1,20) = 14.455, p=0.001) in such a way that the category advantage (how much faster subjects

---

[1] It is important to note that although participants were faster on the between-category trials than on the within-category trials in the no interference condition we cannot draw any strong conclusions about this advantage. It is possible that the psychophysical distance for two chips in the between-category condition was slightly greater than for those in the within-category condition. The critical comparison in the design of this experiment (and in all of the experiments in this paper), is the interaction between interference type and whether or not the color chips in a comparison belong to the same or to different linguistic color categories. That is, we are interested in whether the difference in performance on between-category versus within-category trials changes as a function of verbal interference. Since subjects make identical sets of discriminations on identical stimuli in each interference condition, any non-uniformity in the spacing of the chips in color space would be the same across interference conditions, and so could not produce such an interaction.

were on the between-category trials than the within-category trials) was greater for trials on which the foil differed by 1 Munsell chip rather than 2 (mean category advantage at distance 1=197, SE = 51, mean category advantage at distance 2 = -19, SE = 51).

Critically, the effect of category interacted with the effect of interference (F(1,20) = 7.272, $p<0.05$, figure 5) so that the categorical advantage for between-category trials relative to within-category trials was significantly reduced by verbal interference (mean category advantage with no interference = 167 ms, SE = 44, mean category advantage with verbal interference = 11 ms, SE = 58). There were no other significant interactions.

Because stimuli stayed on the screen until the subjects responded, reaction time was the measure of interest, and accuracy was high. However, we also examined the accuracy data in order to check for any speed-accuracy trade-offs. Of particular interest was whether a similar interaction would be seen between category and interference in the accuracy data as we saw in the reaction-time data. Analysis showed that there was a significant interaction between category and interference in the accuracy data, (F(1,20)=7.7, p = .026). Subjects were less accurate on between-category than on within-category trials in the no interference condition (between = 80%, SE =2%; within = 87.5%, SE = 2%; advantage = 7.5%, SE 2%). This difference was reduced in the verbal condition (between = 82.5%, SE 2%; within = 83%, SE = 2% ; advantage = 0.5%, SE = 2%). A paired t-test showed that this reduction was significant (t(20) = 2.4, $p < 0.03$, all t-tests are 2-tailed). To assess whether this effect on accuracy was related to the effect on reaction time we correlated each subject's category advantage in reaction time with the advantage in accuracy for each interference condition. That is we tested whether the difference in reaction time for between and within category trials was related to the difference in accuracy data within each subject. The results of the correlations showed no significant relations, ruling out the possibility of speed-accuracy tradeoffs (no interference $r(20) = -0.101$, $p>0.5$; verbal interference $r(20)= -0.13$, $p > 0.5$).

3.3 Discussion

Results of Experiment 1 suggest that categorical distinctions that exist in language can shape people's performance even in a very basic perceptual discrimination task. Interfering with subjects' normal ability to verbalize changed the pattern of results in perceptual discrimination across a linguistic color boundary. This was true even though subjects were performing an objective implicit task where all stimuli were simultaneously available on the screen during perceptual judgments. These results support the hypothesis that language plays an online/active role in perceptual judgments and that linguistic information is habitually recruited in the normal course of making simple perceptual decisions.

However, it is possible that the pattern of results obtained in Experiment 1 is not specific to verbal interference per se. While verbal interference diminished the categorical advantage found without interference, it is not clear if verbal interference had this effect specifically because it interfered with some necessary aspects of linguistic processing, or if any secondary task that requires the same amount of effort or attention (but does not tie up linguistic resources) would have had the same effect. In Experiment 2, subjects performed color judgments as before either with no interference, with verbal interference, or with a new spatial interference task in order to test whether the effects of verbal interference found in Experiment 1 were indeed specific to tying up linguistic processes or merely due to general cognitive load.

4. Experiment 2

In Experiment 2, participants performed the same color-judgment task as in Experiment 1. On one third of the trials, participants made color judgments without interference, on one third of the trials while performing a verbal interference task as before, and on a final third while performing a spatial interference task. If the decline in the categorical advantage we observed under verbal interference in Experiment 1 is due specifically to tying up linguistic resources, then a similar decline should not be seen under spatial interference. On the other hand, if the reduction in categorical advantage during verbal interference was due to tying up domain-general resources, then we might expect a similar decline in the categorical advantage during both spatial and verbal interference.

4.1 Methods

4.1.1 Subjects

49 native English speakers participated in the study in exchange for payment.

4.1.2 Design & Materials

This experiment used the same stimuli, equipment, and testing rooms as Experiment 1. Trials were blocked by interference type, (spatial, verbal, or none) with 96 trials in each condition. The order of blocks was random across subjects. The color matching trials were identical to those in Experiment 1 except that only foils that were 1 Munsell chip away from the match were used, largely to simplify the analysis of the data.

For the spatial interference block, subjects were presented with a square 4x4 black and white grid (see Figure 6 for an example). Within the grid pattern, 4 of the 16 squares had been blacked out. Subjects were presented with the grid for 3 seconds and asked to

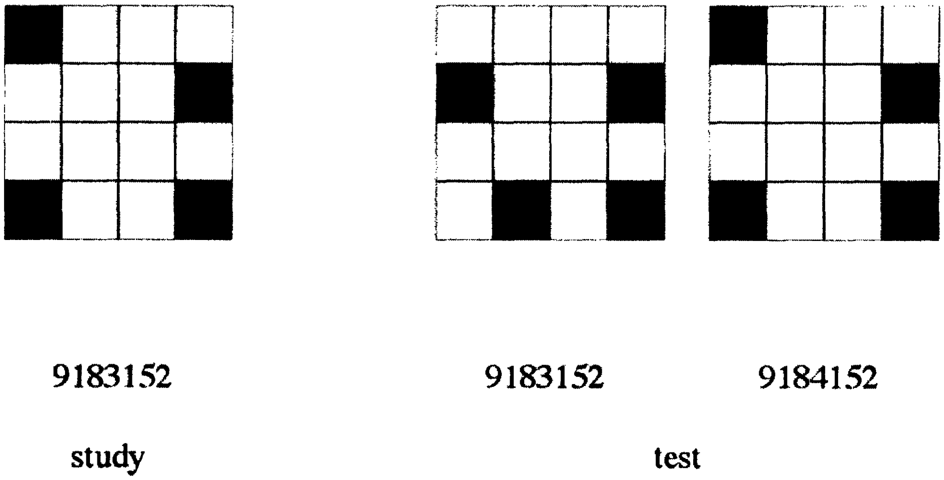


Figure 6. Examples of spatial and verbal interference used in experiments 2 and 3. Lure items at test differed from the target either by 1digit for the verbal items or by the displacement of 1 black square for the spatial items. Pilot studies showed that subjects performed equally well with the two kinds of stimuli.

remember it by maintaining a visual image of the pattern. Subjects then performed 8 color matching trials. Subjects were then shown the original grid and a foil and were asked to indicate which one they saw before. Foils were created by moving a single black square from the sample grid pattern one space in any direction. On verbal interference trials, subjects were presented a number string and asked to remember the number by rehearsing it. Just as with the spatial interference, after 8 color matching trials, subjects were shown the original number string and a foil and were asked to indicate which one they had seen before. Verbal foils were created by randomly changing one of the digits. For the verbal and spatial interference blocks there were 12 numbers and 12 grids to be recalled respectively.

We conducted a pilot experiment with 11 subjects to ensure that the spatial task was at least as demanding as the verbal task with respect to accuracy. On each trial of the experiment subjects were shown either a 4x4 grid or an 8 or 9 digit number for 3 seconds. After a 5 second delay, subjects were asked to select the stimulus they had seen before from 2 alternatives. Foils were made as described above. Performance was numerically slightly better for the 8 digit numbers than the 4x4 grids, but statistically no difference was found (grids: 95% correct, SE = 1%; 8 digits: 96% correct, SE = 1%; 2 tailed $t(10) = 0.94$, $p>0.35$). There was a significant difference in reaction time, with subjects taking longer to choose between the two alternatives for the verbal items (grids: 1643 ms, SE = 173 ms; 8 digits: 2478 ms, SE = 260 ms; $t(10) = 4.5$, $p<.005$). This difference can likely be attributed to the speed with which the target and lure can be processed and compared, but if anything, this is evidence that the verbal task is easier. That is because cost in memory performance due to items forgotten during the test will lead us to overestimate the difficulty of the verbal interference task, making it possible that the verbal interference task is easier than spatial interference task. Since the point of matching the difficulty of the spatial and verbal interference task is to show that any selective effect of verbal interference is not due to the verbal interference using more general memory resources, the possibility that it might be easier than the spatial interference task is not a confound. Based on these pilot data, we decided to use 7- and 8-digit number strings for the verbal interference manipulation in Experiment 2. For 25 of the subjects the number was 8 digits in length and for the remaining 24 the number was 7 digits long. The fact that the spatial task was either as difficult or slightly more difficult than the verbal task provides us with a conservative test of whether the effects of verbal interference in Experiment 1 were due to the interference task tying-up domain-general resources, or linguistic resources in particular. Note that the verbal interference task was changed slightly from the free response task used in the previous experiment in order to make it more similar to the spatial interference task and to match the difficulty level of the two tasks.

4.2.1 Analysis

4 subjects whose mean reaction times were more than 2 standard deviations from the overall mean were removed from the analysis. These same criteria were applied to Experiment 1 (and to all subsequent experiments), but in Experiment 1 no subjects met the exclusion criteria. Incorrect discriminations and trials occurring during blocks where the verbal or spatial items were not recalled were excluded from the analysis. This left on average 83% of each subject's trials (SE 1%). These trials were analyzed using a 3 (interference type: none, spatial, verbal) x 2 (category: between-category, within-category) x 2 (interference length 7 digits or 8 digits) ANOVA with interference and discrimination as a within subjects measure and verbal interference length as a between subjects measure.

4.2.2 Overview of Results

Results are shown in Figure 7. As in Experiment 1, subjects showed a categorical advantage when tested without verbal interference: they were faster on between-category trials than on within-category trials. As before, verbal interference reduced the categorical advantage effect relative to no interference. No such reduction was observed in the spatial interference condition. This result supports the idea that the decrease in categorical advantage seen under conditions of verbal interference results from occupying specifically linguistic resources, and not just from any kind of interference.

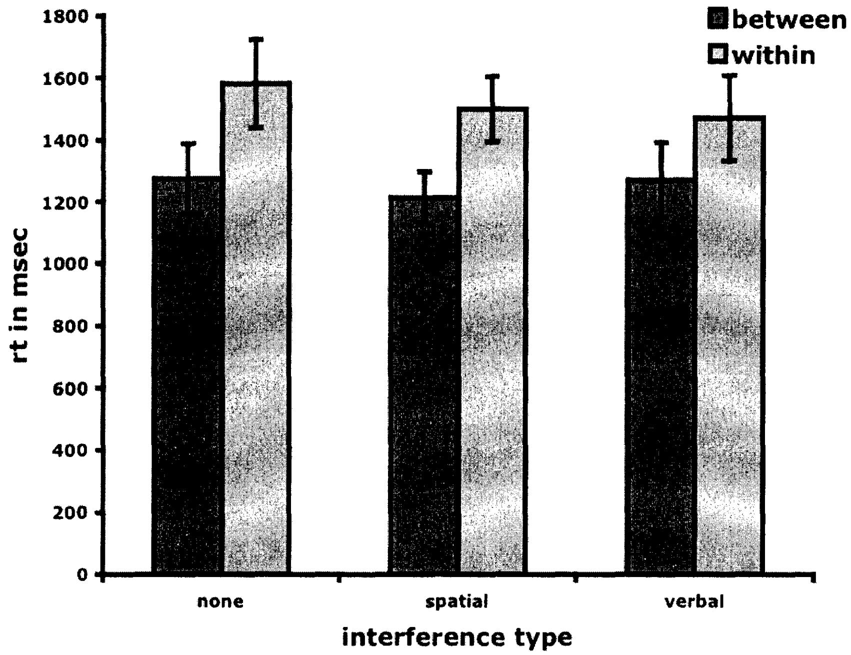


Figure 7. Selective effect of verbal interference on the category advantage found in Experiment 2. The y-axis shows the reaction time in milliseconds. While subjects are significantly faster to do the task on between category trials, that advantage is selectively reduced by verbal interference but not spatial interference relative to no interference. Error bars represent standard error of the mean.

4.2.3 Detailed Analyses

There was a main effect of category (F(1,43)=89.3, $p<0.001$) with between-category trials answered significantly more quickly than within-category trials (mean between = 1283 ms, SE = 87 ; mean within = 1566 ms, SE = 99). There were no other main effects and digit length did not significantly interact with any other factor.

Importantly, there was an interaction of interference type by category, (F (2,43) = 7.1, p = 0.001). Planned t-tests showed that the size of the category advantage was not different between the spatial interference and no-interference conditions (category advantage with spatial interference = 306 ms, SE = 34; no interference = 343 ms, SE = 42 ; t(44) = 0.92 p = 0.36). However, this categorical advantage was significantly reduced under verbal interference (category advantage with verbal interference = 203 ms, SE =

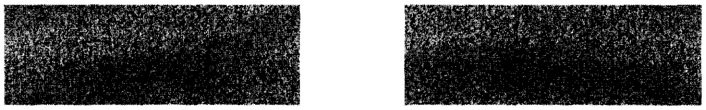
36) relative to both the no interference trials (t(44) = 3.7, $p < 0.001$), and the spatial interference trials (t(44) = 2.9, p = 0.005). These results suggest that the interference-related decline in category advantage in this and the previous experiment was the result of occupying specifically linguistic resources.

While reaction time was the measure of interest, we analyzed the accuracy data to test for any speed accuracy tradeoffs. An ANOVA on the accuracy data showed no significant interaction between category and interference (F(2,43) = 0.77, $p>0.45$). As a further check we again assessed the correlation between the categorical advantage in reaction time and the advantage in accuracy for each subject under each type of interference. As before, none of these correlations approached significance thus ruling out speed-accuracy tradeoffs as an explanation for the pattern of results observed in reaction times (no interference r(43) = .091 p = 0.552; spatial interference r(43) = .147 p = 0.335; verbal interference r(43) = -0.16, p = 0.294).

4.3 Discussion

Results of Experiment 2 replicated and expanded the results of Experiment 1. In Experiment 2, verbal interference reduced the categorical advantage across the blue-green boundary. A matched spatial interference task did not reduce this categorical advantage. This demonstrates that the reduction in category advantage is a result of occupying specifically linguistic resources rather than of a generic increase in cognitive load. These findings further support the hypothesis that language plays an online/active role in even simple and objective perceptual judgments. One difference between the two experiments is that in Experiment 1, verbal interference entirely eliminated the categorical advantage while in Experiment 2 it did not. One possibility is that in experiment 1, subjects were required to type in the 7 digit number rather than simply select from 2 alternatives. This likely increases the memory load (retrieval vs. recognition) and this increased difficulty may account for the difference between the two results.

In Experiments 1 and 2 memory demands were minimized by making all of the color chips simultaneously available on the screen until subjects responded. While the memory demands of such a task are minimal, it is possible that comparing colors even across a small bit of white space still requires some very short-term visual memory. It is possible for example that subjects shifted their gaze or attention between the test patches in the course of doing the task. In Experiment 3 we further simplified the task by removing the white space separating the test patches. Experiment 3 used the same experimental design as Experiment 2 but with all the chips placed directly adjacent to one another, forming a single rectangle. One end of the rectangle was different in color from the other two thirds of the rectangle (Figure 8). Subjects were instructed to identify the side of the rectangle containing the differently colored square patch. To perform this task subjects needed only to be able to detect a color edge, and did not need to make any comparisons across white space



Your job is to say which side the square patch is on.

WITHIN CATEGORY

BETWEEN CATEGORY

Figure 8. Examples of stimuli used in color task for experiment 3.

5. Experiment 3

5.1 Methods

5.1.1 Subjects

45 native English speakers recruited at Stanford and MIT participated in the study in exchange for payment or course credit.

5.1.2 Design & Materials

Experiment 3 used the same stimuli and equipment as Experiments 1 & 2, and the same experimental design as Experiment 2. Because no differences were found between 7- and 8-digit strings in verbal interference in Experiment 2, only the 8-digit strings were used in Experiment 3. The main design difference was that the three squares in the color triad were rearranged to form a single continuous horizontal bar (Figure 8). The foil was located either on the extreme left or right of the bar. On each color discrimination trial subjects were asked to respond whether the different color chip was on the left or the right and to respond by pressing a key on the left or right side of the keyboard.

5.2.1 Analysis

Three subjects were removed for having slow reaction times using the same criteria as in Experiments 1 and 2. Only trials where the correct answer was given and the interference item correctly remembered were analyzed, with an average of 98%, SE =0.005% of each subject's trials left for the analysis. Reaction times were analyzed in a 3 (interference type: none, spatial, verbal) x 2 (discrimination type: between, within) ANOVA.

5.2.2 Overview of Results

As shown in Figure 9, subjects were faster on trials in which the distracter crossed a linguistic boundary than on trials where the distracter came from the same linguistic category. Unlike Experiments 1 and 2, neither verbal nor spatial interference significantly modulated the advantage for between category trials. Detailed statistical analyses are reported below.

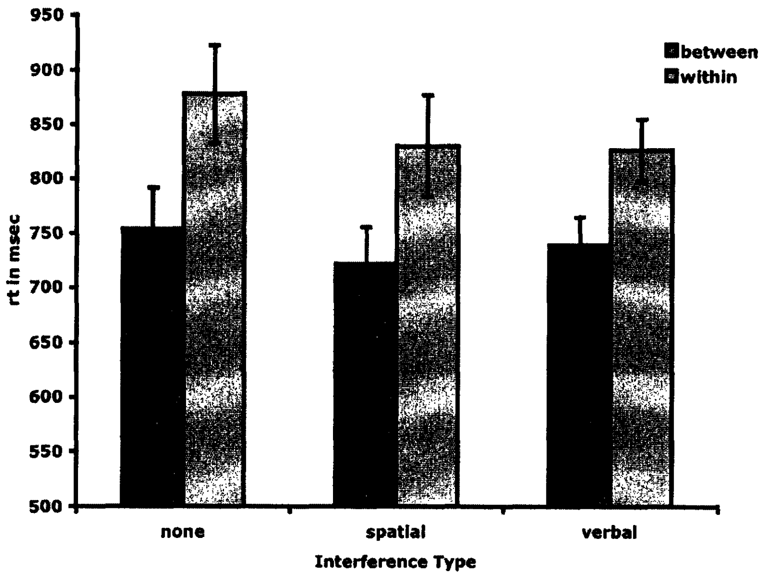


Figure 9. No effect of interference on category advantage in experiment 3. When the spatial arrangement of the stimuli is changed so that all the stimuli are next to each other (as in figure 8) the selective effect of verbal interference on the category advantage disappears. Note that the overall reaction time in this experiment is much faster than in experiments 1 and 2.

5.2.3 Detailed Analyses

There was a significant main effect of category with faster reaction times for between category discriminations (between = 738 ms, SE = 26, within = 844, SE = 32, $F(1,41) = 41, p<0.001$). There were no other significant effects. The category advantage for each of the interference types can be seen in Figure 9. Unlike Experiments 1 and 2, there was no selective effect of the verbal dual task on the category advantage. To verify that the pattern of results differed between experiments 2 and 3, we analyzed the performance using a 2 (experiment: 2, 3) x 3 (interference: none, spatial, verbal) x 2 (category: between, within) ANOVA. The relevant result was that we found evidence for a three way interaction between experiment, interference, and category (F(2, 85) = 2.86, p = 0.06), suggesting that the effect of interference type on category differed between the two experiments.

While overall accuracy for experiment 3 was near ceiling, an ANOVA of the accuracy data revealed a significant interaction between category and interference (F(2,40) = 3.055, p = 0.053) with within-category trials being less accurate than between-category trials during verbal interference (mean between = 99%, SE = 0.2%; mean within = 98.6% with SE 0.3%). One possibility, is that this very small but reliable change in accuracy is masking differences in reaction time performance. However, this would seem to mean that if accuracy were equal subjects should go slower on the within-category trials thereby increasing the category advantage for between-category trials (at least in terms of reaction time) and further mitigating any difference between verbal interference and the other types. As in Experiments 1 and 2, correlations between the category advantage in reaction time and accuracy for each interference condition showed no significant relationships further ruling out the possibility of speed-accuracy tradeoffs (no interference r(41) = 0.023, p = 0.885; spatial interference r(41) = -0.002, p = 0.991; verbal interference r(41) = -0.113, p =0.476).

5.3 Discussion:

Arranging the color chips in a continuous rectangle such that all squares abutted, eliminated the effect of language on color discrimination that we had found in previous experiments. While subjects were still faster on trials where the foil came from a different linguistic color category, this advantage was unaffected by either spatial or verbal interference. These results demonstrate that while language is recruited for some low-level perceptual tasks, not all perceptual tasks reveal effects of the online involvement of language. Identifying the critical differences between tasks that do and don't reveal the involvement of language will provide a good subject for further study. Here we suggest a number of possibilities.

One possibility is that the task used in Experiment 3 is simply too easy and too fast to be affected by language. The reaction times in this task were considerably faster than in the tasks used in Experiments 1 and 2 (Experiment 3 mean = 792 ms, SE = 29 ms; Experiments 1 and 2 combined mean 1392 ms, SE = 69ms). It is possible that linguistic (or other higher-level) knowledge can only affect or assist the processing in a perceptual decision if the task is sufficiently difficult and allows enough time for perceptual information to be passed on to higher-level areas, which can then contribute to the response. This potential explanation makes the prediction that effects of language should be seen in more difficult or longer tasks, and not in very easy or more simple tasks that could be solved long before there is time for perceptual information to reach higher-level areas. One way to answer this question would be to reduce the perceptual difference between the color patches while maintaining between and within category distinctions, thus forcing subjects to perform more slowly. If the overall duration of the task is critical in bringing language to bear on perceptual judgments, the selective disruption of the categorical advantage by verbal interference should reappear when subjects need more time to determine the answer.

However, more recent work using the interference paradigm found here in conjunction with visual search has found effects with mean reaction time of less than 500 ms (Gilbert, Regier, Kay, & Ivry, 2006) suggesting that online effects of language can be found even in tasks with faster reaction times than those in Experiment 3. Gilbert et al showed that verbal interference selectively diminished the advantage subjects had for finding targets that were from a different linguistic category, but only when these targets appeared in the right visual field. This opens up the possibility that if Experiment 3 were repeated but visual field controlled, an effect might be found for targets in the right visual field but not the left.

Another potential source of difference between the triad tasks used in Experiments 1 and 2 and the rectangle task used in Experiment 3 (as well as the visual search task used by Gilbert et al) might be the memory requirements of the triad task. While all stimuli were available simultaneously on the screen, they were separated by narrow bands of white space. It is possible that subjects attended to the color squares in the triad one at a time, and had to remember the colors across attentional shifts or saccades in order to compare the patches (Newhall, Burnham, & Clark, 1957). In Experiment 3, the correct answer can be determined merely by finding a color edge (i.e. without doing a comparison to another item in memory). It is possible that the minimal requirement of memory across attentional shifts or saccades in Experiments 1 and 2 was enough to recruit the linguistic system to help as a secondary/more compact code. Overall it is surprising that even the rather simple perceptual matching task used in Experiments 1 and 2 would show effects of language. Tasks with similar or greater attentional/memory requirements are routinely used in vision science to determine perceptual distances and just-noticeable differences. More generally, comparing stimuli across saccades or to items stored in memory are tasks that represent a great proportion of natural visual behavior, and arguably more closely resemble natural visual behavior than comparing two abutting colors.

6 General Discussion:

Does language influence English speakers' judgments about color? Two of the experiments presented in this paper demonstrate that language plays an active, online role in even very simple perceptual tasks. We find that English speakers tested across the blue/green boundary in English show a categorical distortion in performing a simple perceptual matching task, and this categorical distortion can be reduced by a verbal, but not a spatial, dual task. These results demonstrate that categories in language can affect performance on basic perceptual color discrimination tasks. Further, they show that the effect of language is on-line, since it is disrupted by verbal interference. Results from a third experiment suggest that not all perceptual tasks are affected by language. Understanding why linguistic categories may play a role in some perceptual decisions and not others will help further clarify the mechanisms through which language can affect perceptual performance.

Some hints about the mechanisms through which linguistic categories can influence perceptual performance can be gleamed from the current set of results. It appears that the influence of linguistic categories on color judgments is not limited to tasks that involve remembering colors across a significant delay. In our task, subjects showed language-consistent distortions in perceptual performance even though all colors were in plain view at the time of the perceptual decision. Further, language-consistent distortions in color judgments are not limited to ambiguous or subjective judgments where subjects may explicitly adopt a language-consistent strategy as a guess at what the experimenter wants them to do. In our task subjects showed language-consistent distortions in perceptual performance while making objective judgments in an unambiguous perceptual discrimination task with a clear correct answer.

Results from the verbal interference manipulation provide further hints about the mechanism through which language shapes perceptual performance in these tasks. One way that language-specific distortions in perceptual performance could arise would be if low level visual processors tuned to some particular discriminations show long-term improvements in precision, while processors tuned to other discriminations become less precise or remain unchanged (Goldstone, 1994). Very specific improvements in perceptual performance are widely observed in the perceptual learning literature, and are often thought to reflect changes in the synaptic connections in early sensory processing areas (Karni, 1996). The design of the current study does not allow us to test for this possibility. However, we do find that a simple task manipulation – asking subjects to remember a digit string - reduces language-specific distortions in discrimination. If the language-specific distortions in perceptual discrimination had been entirely a product of a permanent change in perceptual processors, then temporarily disabling access to linguistic representations with verbal interference should not have changed the pattern in perceptual performance.

Our results suggest that language-specific distortions in perceptual performance arise at least in part as a function of the interaction of lower-level perceptual processing and higher-level knowledge systems (e.g., language) online, in the process of arriving at perceptual decisions. The exact nature of this interaction cannot be determined from these data. It could be that information from linguistic systems directly influences the processing in primary perceptual areas through feedback connections, or it could be that a later decision mechanism combines inputs from these two processing streams. In either case, it appears that language-specific categorical representations play an on-line role in simple perceptual tasks that one would tend to think of as being primarily sensory. Language-specific representations appear to be brought online spontaneously during even rather simple perceptual discriminations. The result is that people show different patterns in perceptual discrimination performance when tested under normal viewing conditions versus when normal access to language-specific representations is disrupted (as in the verbal interference condition).

Since linguistic categories differ across languages, the precise regions in color-space where linguistic distinctions come into play should also differ across speakers of different languages. Indeed, we have found that these kinds of categorical distortions vary between Russian and English speakers across an obligatory language boundary in Russian, that does not exist in English (Winawer et al., 2007; Witthoft, 2003). Unlike English, Russian does not have a single color word for blue, and instead makes an obligatory distinction between lighter blues ('goluboy') and darker blues ('siniy'). There is no single generic word for "blue" in Russian that can be used to describe all of the colors called "blue" in English. Using the same triad color-matching and interference tasks as in Experiments 1 and 2, we found a categorical advantage in Russian speakers across the Russian goluboy/siniy border (Winawer et al., 2007; Witthoft, 2003). This categorical advantage was reduced by verbal interference but not by spatial interference. The same effects did not hold for English speakers. Verbal interference differentially affected only Russian speakers' performance within this blue range, and not English speakers performance. These findings again demonstrate that the particular color categories that exist in one's language play an active, on-line role in rather simple perceptual color decisions.

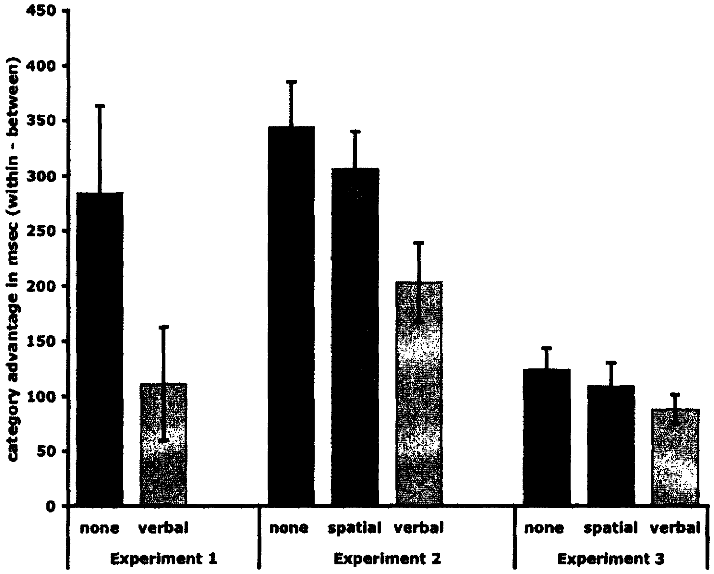


Figure 10. Effect of interference on category advantage (rt on within-category trials – rt on between-category trials) across experiments. In experiments 1 and 2, verbal interference selectively reduces the category advantage. Since only differences of 1 Munsell step were used in experiments 2 and 3, only data from distance 1 in experiment 1 are shown. This effect of verbal interference is not observed in experiment 3.

Conclusion:

This paper presents a series of experiments showing that language can play an online role in what seem to be very simple and objective visual tasks. In Experiments 1 and 2, an effect of verbal interference was found on a color discrimination task even though all color stimuli were presented on the screen at the same time, minimizing the possibility of memory-based distortions (Figure 10). Moreover, a spatial interference task matched for difficulty did not affect performance suggesting that the differences found under verbal interference are specifically due to a change in the availability of linguistic resources. These findings shed some light on the mechanisms through which categories found in language can play a role in perceptual decisions. It appears that linguistic information is habitually recruited in the course of normal perceptual decisions, even in tasks with minimal memory demands, and with an objective correct answer that can be determined in the absence of any linguistic knowledge. While we find evidence for a role of language in some perceptual tasks, this role may be diminished or eliminated, depending on the spatial arrangement of the stimuli (Experiment 3). Future work aimed at understanding why language plays a role in some perceptual tasks and not others will help further clarify the mechanisms through which linguistic categories can affect perceptual performance.

Chapter 4: Adaptation and Face Recognition

Introduction

Perceptual adaptation and aftereffects have long interested psychologists both for the clues they provide about what kinds of features (and where) might be coded in the visual pathway and because they demonstrate the ways in which the visual system adjusts itself in response to experience(Held, 1980; Barlow, 1990). While many well known aftereffects have been described for ‘simple’ visual features like color and the orientation of lines (Gibson & Radner, 1937), more recent work has found that adaptation and aftereffects are found with ‘complex’ stimuli such as shapes and faces (Suzuki & Cavanagh, 1998; Fang & He 2005; Webster & MacLin 1999). Work on face adaptation in particular has shown that adaptation can bias different kinds of judgments about faces including gender, normality, race, and emotion (Webster et al, 2004; Rhodes et al, 2004; Rhodes et al, 2003). It has also been shown that these adaptation effects are somewhat robust to changes in position, orientation, and size, and persist across short delays between adaptation and test (Yamashita et al 2005; Zhao & Chubb, 2001; Leopold et al, 2005). These invariances are thought to be typical of neurons higher up in the visual hierarchy and the persistence of adaptation under these conditions is an indication of a relatively high level locus of adaptation.

Most of these studies induce their effects in the laboratory using similar paradigms. For example, Mike Webster (1999) demonstrated adaptation to faces that had been distorted by either compressing or expanding a region between the eyes. Subjects were first shown faces spanning the continuum from expanded to contracted and asked to judge how normal the faces looked. The subjects then adapted to either a contracted or expanded face by staring at it for some time and then were again presented with a test face taken from the continuum of faces and asked to judge its normality. In between judgments of normality subjects again stared at the distorted face to maintain the adapted state (what is often referred to as top-up adaptation). Webster found that when subjects adapted to an expanded face, subsequent judgments were biased in such a way that slightly expanded faces looked normal and undistorted faces looked compressed. The effect is quite powerful and given sufficient adaptation can be seen by an observer on a single trial which is why there are few objections to the fact that the task itself is a subjective one.

Similarly, Webster has provided a nice demonstration of identity driven aftereffects using a morph between John Kerry and George Bush (demo available from http://viperlib.york.ac.uk). A picture of Bush and a picture of Kerry appear side by side and subjects adapt by fixating a point between the two pictures for half a minute or longer. After this period of adaptation both images are replaced with test faces that are the average between Bush and Kerry. Although the test faces are identical, the face in the location where Kerry was looks more Bush-like and the face where Bush was looks more Kerry-like. While the after-effect is highly visible, with the two test faces looking very different, it is not clear whether the Kerry adapted test looks more like Bush, or just less like Kerry, making Bush the better of the two alternatives. Moreover, is the adaptation effect specific enough that if you did not know Bush was the other endpoint and were unable to recognize him in the averaged face, would adapting to Kerry make Bush visible?

One interpretation of the above data is that the viewer's representation of what is normal shifts towards recent experience and that judgments of normality (or gender, ethnicity, or emotion depending on the task) are really comparisons to a moving neutral point. So, when adapting to an expanded face, the neutral point begins to move towards that face and subsequent testing with what used to be normal will appear contracted. This kind of coding scheme would increase sensitivity to small changes among stimuli that are highly similar and could serve as the basis for longer term learning of face categories (Barlow 1990, Webster 2004, Rhodes et al, 2005).

Along these lines, adaptation has also been used to assess the proposal that face identities are represented using a norm or prototype referenced face space (Leopold et al 2001; Rhodes et al 2006; Valentine 1991, Tsao & Freiwald, 2006). Two versions of the face space hypothesis have been put forward, an exemplar based account, and a prototype referenced account (Valentine, 2001; Lewis 2004). In both, each face occupies a position in a multidimensional space where each dimension is some feature. The features and the dimensionality of the space are unknown but are the focus of much research. The two accounts differ in how they account for various aspects of face processing, such as similarity, identification, and caricature effects (Valentine, 2001). In the exemplar account, identity and similarity are computed using Euclidean distances between the faces being compared. Each face occupies a region in the space, and all faces which produce values inside that region are identified as the same. In the prototype account, locations are computed using the direction of the vector that runs from the center of the space (which corresponds to the average face) to a particular exemplar. Faces which lie on that vector are considered the same face. Faces which lie beyond an exemplar on the line originating from the average face are caricatures (Leopold et al, 2001).

Several recent papers using a constructed face space and adaptation have argued for the prototype account. In one, Leopold et al (2001) constructed a space using a large number of 3 dimensional faces scanned from human subjects (Blanz & Vetter, 1999). Having a space allowed them to do a sophisticated experiment showing that when adapted to a particular face A, subjects were likely to judge the prototype face as having the identity of anti-A (that is the face opposite A in the face space). They further showed that this shift in judgments following adaptation was specific to the trajectory connecting a face and its anti-face through the neutral point or prototype. So for example, if subjects were adapted to A and tested with a face that was 10% B, subjects still labeled the test face as B (see figure 1). This is not necessarily to be expected, since in the face space model, adaptation to A should shift the 10%B face towards Anti-A. The authors conjectured that given the briefness of the aftereffect the identity strength of B overwhelms the aftereffect. In their view, this does not happen with the prototype which has no identity strength at all as it is the average of all the identities and therefore the aftereffect is largely selective for a single trajectory through the face space.

One drawback of the Leopold et al (2001) study is that subjects required extensive training not only on the endpoints but also on faces along the trajectories. That is subjects spent hundreds of trials learning to identify all the faces on a trajectory from A to the prototype as A (though not the prototype itself) presented for 200 ms.. This undercuts the model somewhat, since subjects do not automatically perceive faces along the line from a particular exemplar to the prototype (referred to as the identity line) as having the same identity. Moreover, the lack of identity of the prototype might be a result of the way they have constructed the space and trained the subjects since it lies at the intersection of several trajectories and therefore is a mixture of several identities. Another way of putting this is to say that subjects have learned the structure of the stimulus set. Indeed this seems likely since the prototype referenced effects found by Rhodes et al (2006) and Leopold et al (2001) were found despite the fact that the spaces were constructed from different sets of faces (200 male for the former, and 200 male and 200 female for the later). Similarly, single unit data from monkeys in which neurons showed increased responses as human faces deviated from the norm space was acknowledged by the authors to be a likely artifact of training on the stimulus set (Leopold et al, 2006). In that study, one monkey was trained with one face space and the other monkey with another, but equivalent results were reported for both. This fast learning may seem surprising, but other research has shown that humans and monkeys are good at quickly extracting the underlying structure of novel objects that have been created using a parameterized space (Cutzu & Edelman, 1996; Op de Beeck et al, 2001)

It is also difficult to assess the finding that adaptation did not produce aftereffects on morphs which were not on the identity line. For example if subjects had also been trained on a trajectory that went from face A through say the 20%B face to Anti-A' (see figure 1), it is possible that adaptation would be found along that line as well. Since subjects had not been trained on face Anti-A' (corresponding to the ? in the figure), there is no response they could give. Again, finding that adaptation produces aftereffects on trajectories which do not pass through the norm does not necessarily contradict the prototype model, but finding adaptation specific to the trajectory which passes through the average face is critical to discriminating the exemplar and prototype models (Tsao & Freiwald, 2006).

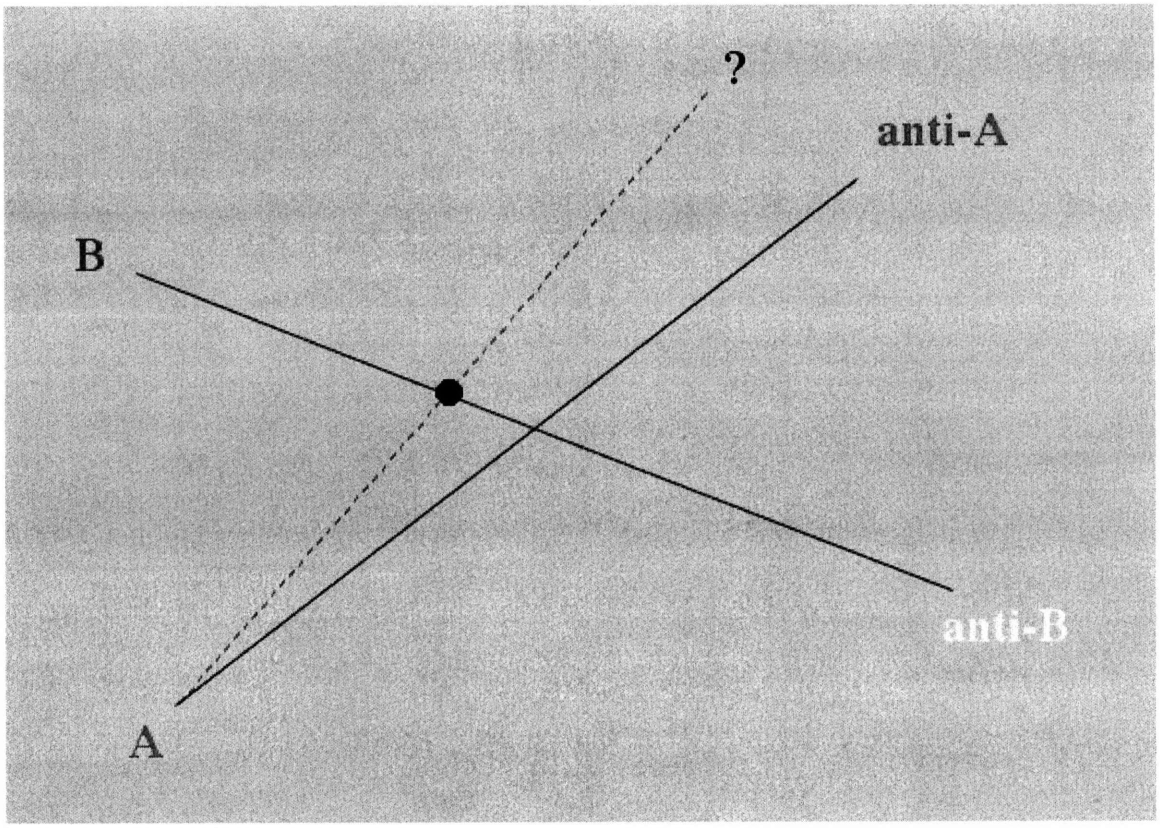


Figure 1: Schematic of Leopold et al 2001 experiment. Solid lines represent morph trajectories between faces and their opposites. The prototype face lies at the center of the space where the two lines intersect. The black dot represents the 10% B face which subjects were trained to classify as B. Although Leopold et al found that adaptation to A did not bias subjects to perceive the 10%B face as anti A, they did not train subjects on the identity that would lie along the dashed line stretching from A to ?.

This paper presents a series of experiments designed to address the issues raised above. We show using an objective method that adaptation can selectively improve subjects' ability to recognize famous faces in morphs made with the adapting stimulus. This effect is obtained without informing the subjects of the identities they will be asked to recognize during the experiment thus removing any possible effects of training or response bias. This provides valuable evidence that the visible distortions in identity aftereffects are specific enough to enhance recognition performance. Furthermore, adaptation effects are obtained on multiple trajectories simultaneously showing that the prototype face and the trajectories which pass through it may not be special, showing that other models of face representation such as the exemplar account may be able to explain identity adaptation effects. Finally, we show some evidence that the effect obtains across transformations in viewpoint, but not up-down orientation consistent with many findings in the face perception literature (Rhodes et al, 2004; but see Leopold 2001).

Experiment 1

In experiment 1, subjects tried to identify faces taken from morphs between famous faces and two unknown faces.

Stimuli and Methods

Ten grey-scale pictures of faces (8 famous and 2 unknown) were used to construct the stimuli in the experiment. Each image was scaled to the same size and then masked using a circle approximately 5 inches in diameter to eliminate external features like the ears and hairline. Using Morph 2.1 software, four of the famous faces (Bill Clinton, Tom Cruise, Christian Slater, and Sylvester Stallone) were morphed with an unknown face. The other four famous faces (Ben Affleck, Bill Cosby, Dustin Hoffman, and Elvis Presley) were morphed with a different unknown face. Each morph line (the set of faces including the endpoints and the graded transformations) was created by hand positioning corresponding points and lines on each pair of endpoint faces. The software then generated 58 intervening images between the endpoints. Counting the unknown face as 1, the 24th, 30th, 36th, and 40th faces were chosen from each morph for use in the experiment. For each morph, these are referred to as the 40, 50, 60 and 67% famous faces respectively. The unknown faces and the famous faces are considered 0% and 100% famous. Examples of the stimuli can be seen in figure 2.

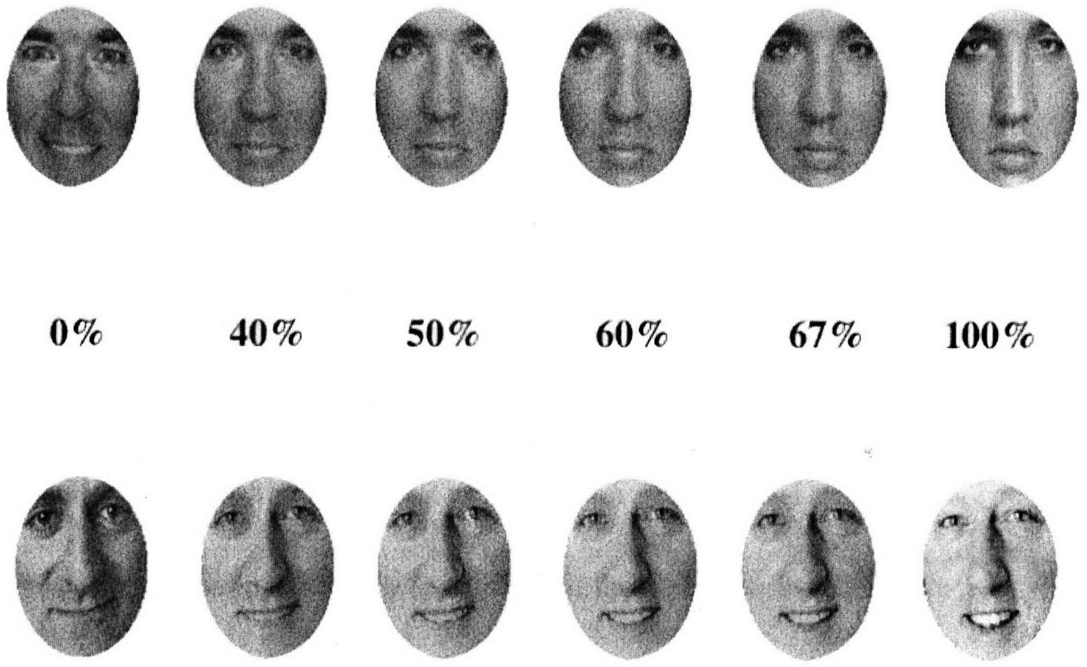


Figure 2: Examples of faces taken from two morphs and used in the experiment. The top row shows faces from Face B to Elvis Presley while the bottom row shows faces in morph from Face A to Bill Clinton.

Stimulus presentation and response collection was done using VisionShell software on Macintosh iMac computers. All subjects were recruited at Stanford and tested in quiet darkened rooms and received either 5 dollars payment or credit in an introductory psychology course.

40 subjects participated in the experiment. 10 subjects adapted to face A during the experiment, another 10 adapted to face B, and the final 20 did not undergo any adaptation. The test stimuli for all the subjects were identical. Each version of the

experiment had 6 blocks with the percent famous face increasing across the blocks. In the first block, subjects were told that they would be shown a series of faces each of which resembled a famous person, and that they should try to guess who it was. The subjects were then shown the 8 40% famous faces in random order and asked to guess what famous person they resembled if they could. At no time before or during the experiment were subjects told the identity of the famous faces used to make the morphs. Subjects generally were unable to identify the famous face at this level, with only 21 correct identifications overall in this part of the experiment (6.6% of tests).

For the subjects in the adapting conditions, each of the next 4 blocks began with 55 seconds of adaptation to one of the two unknown faces. They were then presented with one of the 40% faces for 1 second and asked what famous person they thought it resembled. Subjects were encouraged to guess even if they were unsure. This was followed by 5 seconds of top-up adaptation and then another 40% face until all the 40% faces had been shown. After the 40% faces, subjects did the 50%, then the 60% and then the 67% faces. In the last block, subjects were shown the 100% famous faces and asked to identify them. The 100% faces remained onscreen until subjects made a response. Subjects in the baseline (no adapt) condition had the identical sequence of test stimuli but no intervening adapting faces. Subjects did not fixate during the adaptation or test phases and were instructed to freely view the images.

Our prediction is that subjects should perform better at recognizing famous faces that have been morphed with the stimulus they are adapting to. In other words, less of a famous face needs to be present if it is part of a morph with the adapting stimulus. On the other hand, stimuli that have been morphed with the non-adapting stimulus should not show any particular effect. Since subjects do not know the identities of the famous faces used in the experiment, any information they have must be gotten from the stimuli rather than changes in response strategy (bias). Furthermore, since the test stimuli are the same for all subjects regardless of the adaptation condition, any differences must be attributable to the effect of adaptation.

Analysis and Results

For each subject, faces were excluded from the analysis either if they were recognized at the first exposure of 40% prior to any adaptation or if they were not recognized in the final block at 100%. Data from each of the subjects who were adapted was divided into two groups dependent on whether the test face had been morphed with the unknown face the subject was adapting to, or the one the subject never saw. Test faces from morphs using the adapting stimulus as an endpoint are referred to as 'same', and those using the other face were labeled 'different'. Thus, if a subject adapted to face A, the morphs made with face A were considered 'same' and the morphs made with face B 'different'. The situation is reversed for those subjects adapting to face B. Within each of these groups, the number of correctly recognized faces from a level of morph was divided by the number of faces the subject recognized at the 100% level. This yielded a normalized percent correct score for each level of morph.

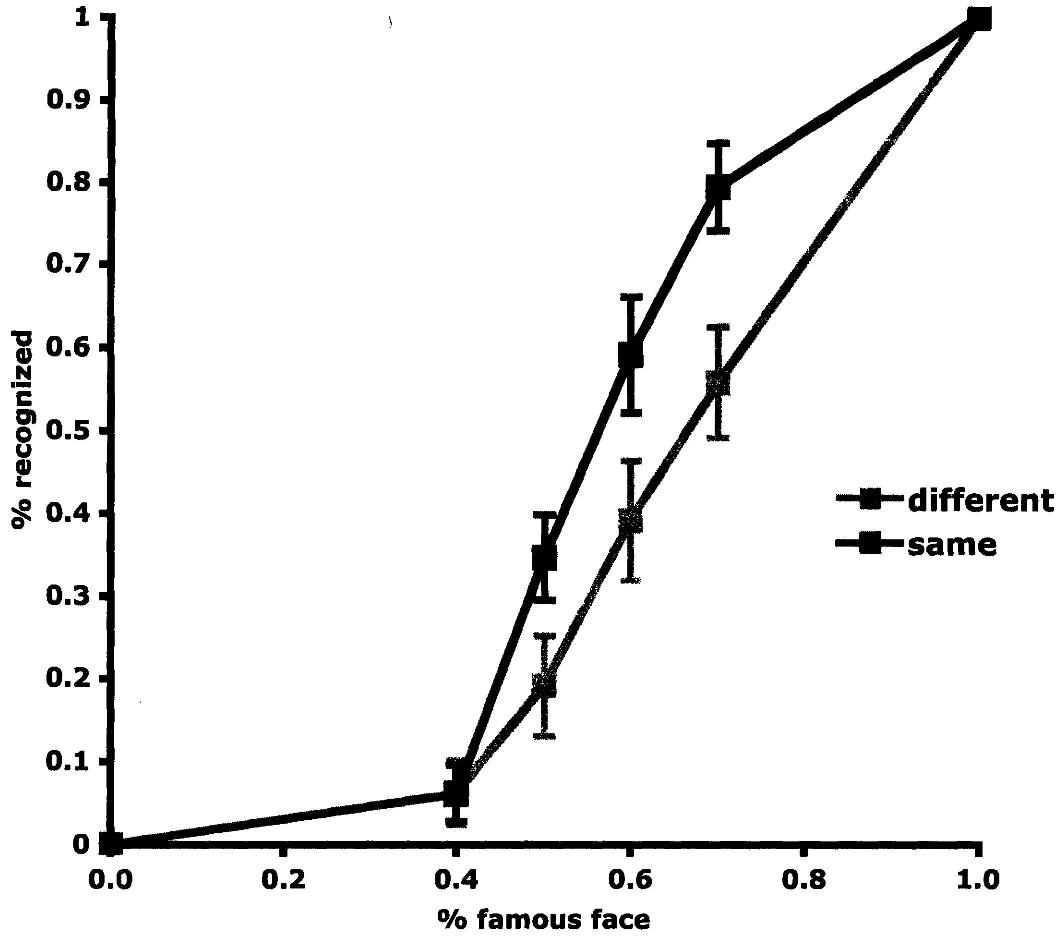


Figure 3. Recognition performance of subjects in experiment 1. Error bars represent standard error of the subject means. Note that stimuli were presented in blocks with each block increasing the percentage of famous face. So the data for the 40% faces on the x axis come from block 2, the 50% from block 3 and so on. Unpaired t-tests show that recognition performance for same items was significantly better for the 50, 60, and 70% test faces. Points represent means across subjects and error bars are standard error of those means.

Overall, subjects did fairly well at recognizing the faces, on average correctly identifying 72% at some point during the experiment (mean number recognized = 5.75 faces out of 8, SEM = 0.32). Every subject recognized at least 3 faces. Figure 3 shows the average data from the 20 subjects who adapted to one of the two unknown faces in experiment 1. The graph shows that faces that were morphed with the adapting stimulus (same) were better recognized than those morphed with the unknown face not adapted to (different). Another way of putting it is that less of the famous face is required to be present for successful recognition in the same condition relative to the different condition.

To quantify the result, data was analyzed in two ways. In the first analysis, performance in the same and different conditions was compared at each level of test face using an unpaired t-test on the subject means. The tests are unpaired since the stimuli labeled same and different vary across subjects. Adaptation produced a significant advantage in recognition performance for the faces that had been morphed with the adapting stimulus at the 50, 60, and 70% levels of test (50%: $t(41) = 2.01$, $p < 0.05$, mean same 35% SE = 5%, mean different 20% SE = 6%; 60%: $t(41) = 1.84$, $p < 0.05$, mean same 59% SE = 7%, mean different 36% SE = 7%; 70%: $t(19) = 3.07$, $p < 0.005$, mean same 79% SE = 5%, mean different 54% SE = 7%). Performance on the 40% tests did not differ as a function of adaptation.

As a secondary analysis a logistic function was fit to the each subject's data (Palmer, Huk, & Shadlen, 2005). The logistic had two parameters, one representing the slope, and one representing the shift of the curve horizontally. Betas representing the horizontal shift in the curve for each subject were determined independently for the same, different, and baseline conditions. Curves that are shifted farther to the left represent better recognition performance (see figure 4). These parameter estimates were then entered into a 1x3 (adapting conditions: same, different, baseline) ANOVA. The results showed a main effect of adapting condition ($F(57,2) = 4.687$, $p < 0.05$. Post hoc t-tests conducted using Tukey's Least Significant Difference procedure showed that a significant shift leftward for the same adaptation relative to the different condition (p = 0.003). The baseline condition was intermediate between the two adapting conditions but did not significantly differ from either (p = 0.1 compared to same and p = 0.14 compared to different).

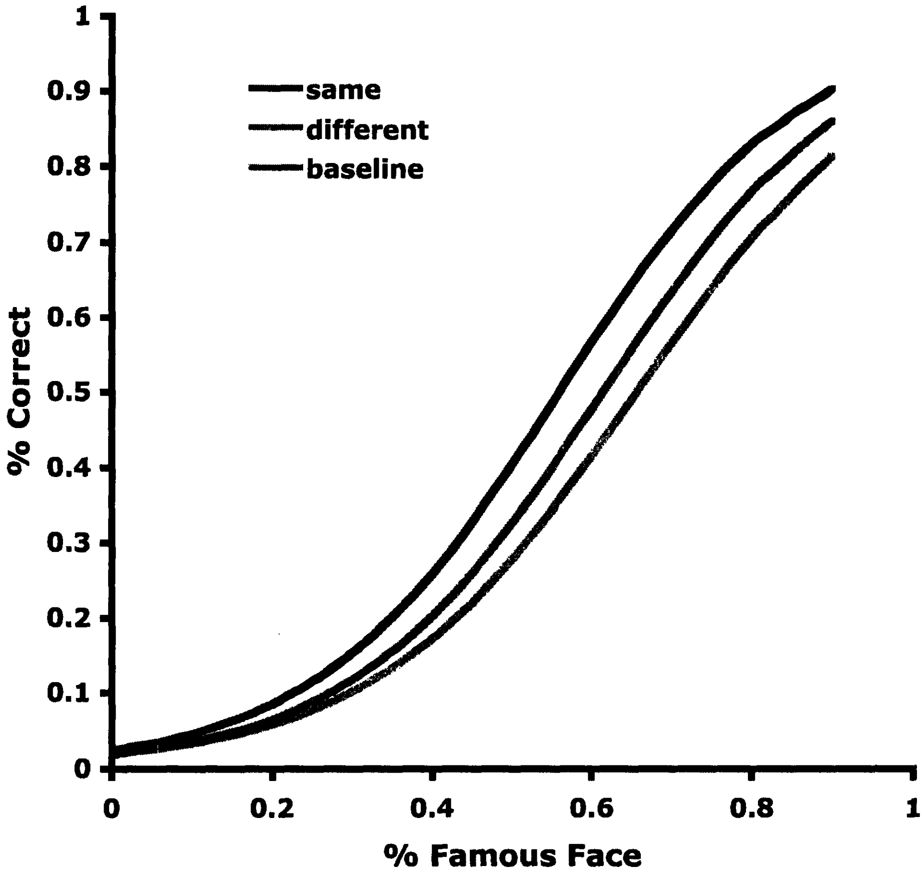


Figure 4: Fitted logistic functions for subjects in experiment 1. Recognition performance is significantly better for faces morphed with the adapting stimulus (purple line) than for those not morphed with the adapting stimulus (green line). Performance in the baseline condition (grey line) where subjects did not undergo any adaptation is intermediate.

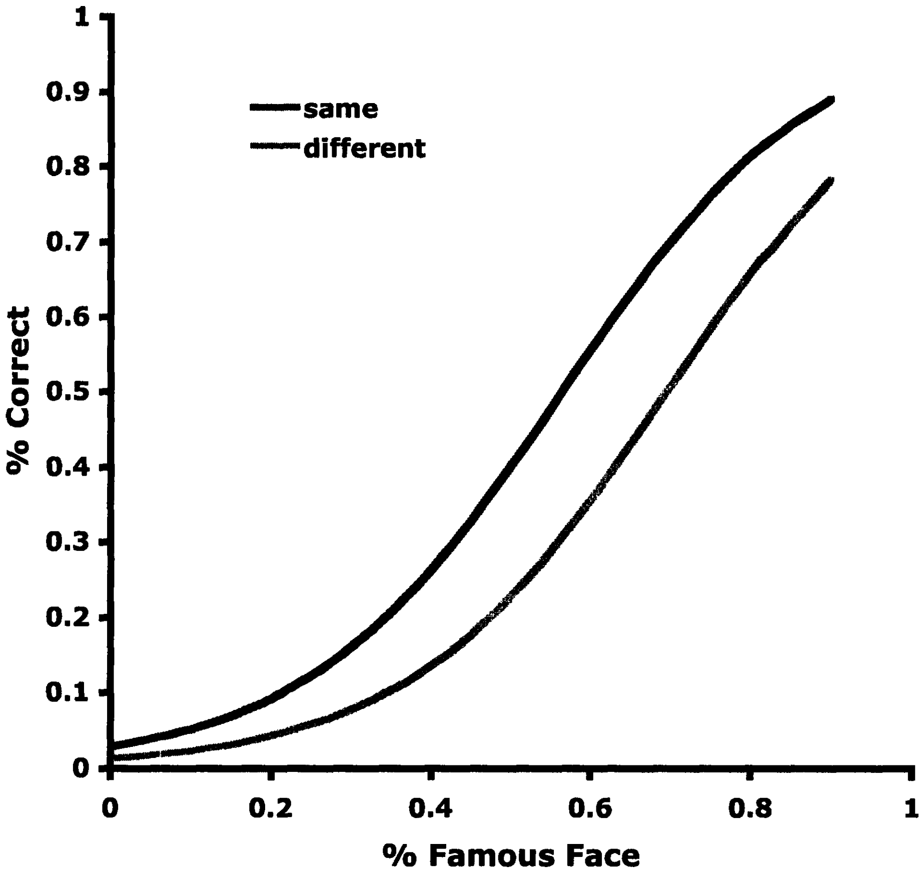


Figure 5. Logistic functions fitted to item data from experiment 1. Subjects required less famous face to be present in the morph in order to recognize it when the adapting stimulus was used in making the morph.

While the preceding finding demonstrates that adaptation can improve recognition performance, the data can also be analyzed by items. To analyze the items two logistic functions were fitted simultaneously for the same and different performance on each item and a set of betas representing the difference in horizontal shift between the two curves was generated (1 beta for each item). These betas were t-tested against zero and showed that better recognition was seen when items were morphed with the adapting stimulus (t(7) = 3.7 , 0.0082). The fitted curves are depicted in figure 5. This way of looking at the data highlights the fact that the improvement in recognition performance cannot be due to some items being easier to recognize than others. The item analysis shows that recognition of a particular face is modulated in the predicted fashion by adaptation.

Experiment 2

Experiment 1 showed that adaptation to an unknown face decreased the amount of a famous face needed in an image in order to recognize it. This effect was specific to morphs that used the same unknown face as an endpoint as an adapting stimulus. The effect of adaptation depended on the adapting stimulus thus ruling out the possibility that the improved recognition was simply due to some of the faces being more recognizable than others. The effect was also present in an item analysis, with 7 of the 8 items showing recognition performance consistent with the adaptation effect. To test the generality of this effect, experiment 2 replicates experiment 1 using 10 new faces (8 famous and 2 unknown).

Methods

The stimuli and design of experiment 2 were identical to experiment 1, except that 8 new famous faces and 2 new unknown faces were used (hereafter referred to as Faces C and D). To increase the likelihood that subjects would be able to recognize the 100% famous faces a separate group of subjects was asked to identify a number of famous faces that had been scaled and masked in the same fashion as they would appear in the experiment. Arnold Schwarzenegger, Elijah Wood, Nicholas Cage, Matt Damon, John Travolta, Harrison Ford, Brad Pitt, and Johnny Depp were recognized the most often were chosen and used in the experiment. One other minor difference was that instead of the 67% morph the 70% morph was used.

34 subjects recruited from the Stanford undergraduate population participated in return for course credit. 11 subjects adapted to face C during the experiment, 11 adapted to face D, and the remaining 12 did not adapt. As in experiment 1 subjects were shown the test faces in blocks with the percentage of famous face in each image increasing across the blocks.

Analysis and Results

Analysis proceeded exactly as in experiment 1 with faces recognized on the first presentation (40%) prior to adaptation and those faces not recognized at 100% discarded from the analysis. Overall recognition performance in the adapting conditions was again good, with subjects recognizing 74% of the faces on average at some point during the experiment (mean number recognized = 5.95, SEM = 0.36).

The results were similar to those of experiment 1. Unpaired one tailed t-tests conducted on each level of the test faces again revealed a effect of adaptation on recognition performance, with significantly improved performance for same relative to different tests on the 50 and 60 and 70% test faces (50%: $t(39) = 2.69$, $p < 0.05$, mean same 35% SE = 5%, mean different 11% SE = 7%; 60%: $t(39) = 3.08$, $p < 0.005$, mean same 61% SE = 8%, mean different 29% SE = 7%; 70%: $t(39) = 2.11$, $p < 0.05$, mean same 70% SE = 8%, mean different 45% SE = 9%). As in experiment 1, there was no difference between the same and different conditions for the 40% tests (Figure 6).

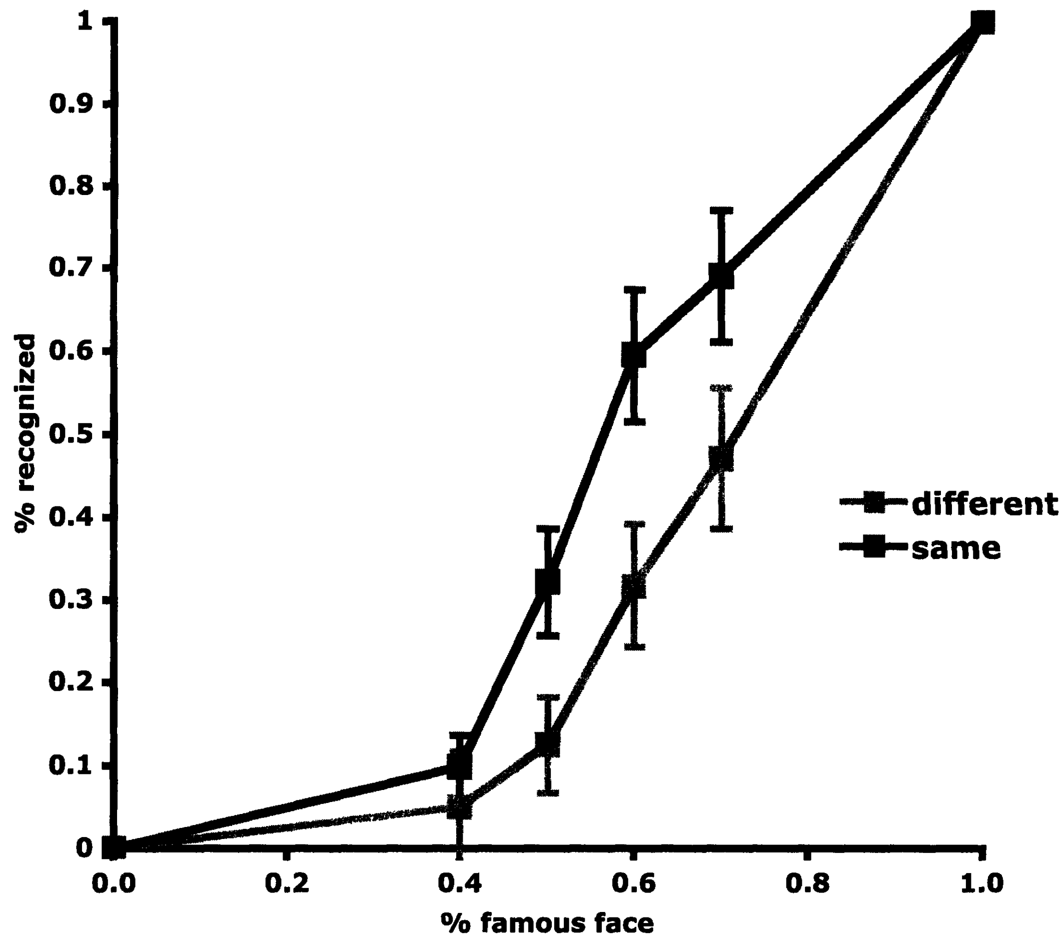


Figure 6: Effect of adaptation on recognition in experiment 2. Purple line represents recognition performance on faces that were morphed with the adapting stimulus. Unpaired t-tests showed significant improvements on recognition for the 50 and 60 and 70% test faces.

Logistic functions were separately fitted to data taken from the same, different, and baseline conditions. Estimates of the parameters measuring the horizontal shift of the curve were entered into a 1x3 ANOVA. Results showed a significant main effect of adaptation condition (F(49,2) = 5.945, p = 0.005). Post-Hoc contrasts showed that again stimuli in the same condition were better recognized than those in the different (p = 0.002). However, in this case, same performance was also better than baseline (p=0.036) while there was no difference between baseline and different (p>.4). This supports the idea that adaptation is boosting performance on the 'same' faces rather than disrupting performance on the 'different' faces.

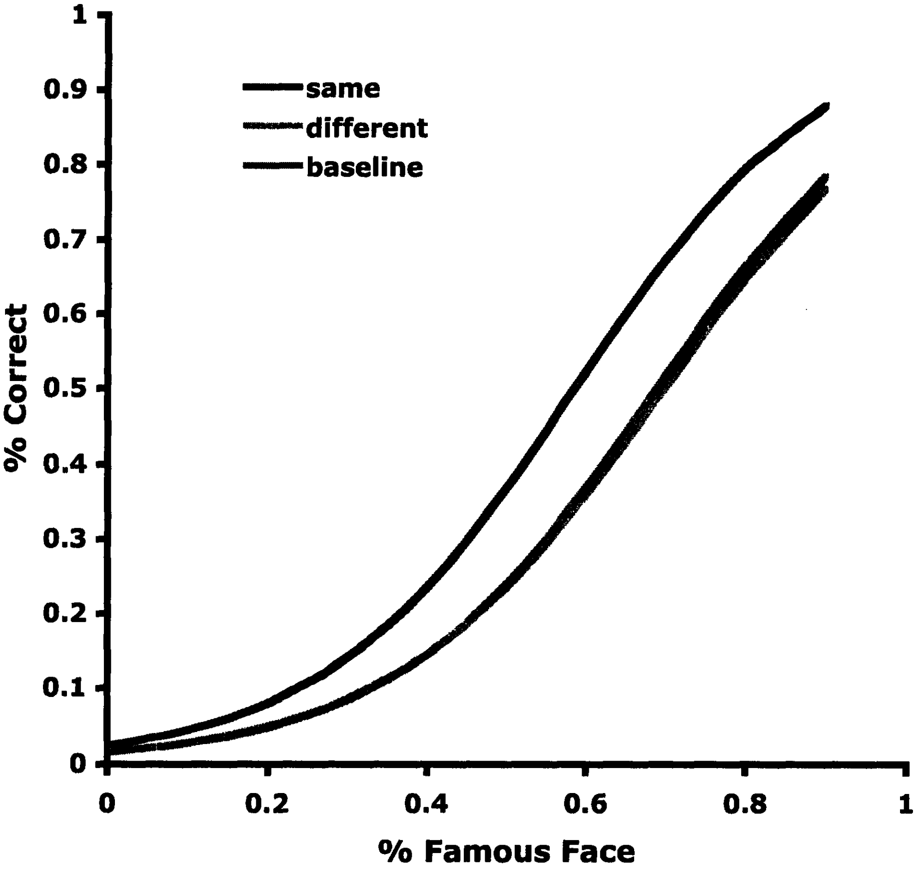


Figure 7. Logistic functions fitted to subject data from experiment 2. As in experiment 1 subjects recognition performance was better when they guessed the identities of faces from a morph that were generated using the adapting stimulus as an endpoint.. However, in this experiment, performance in the baseline condition is indistinguishable from the different condition.

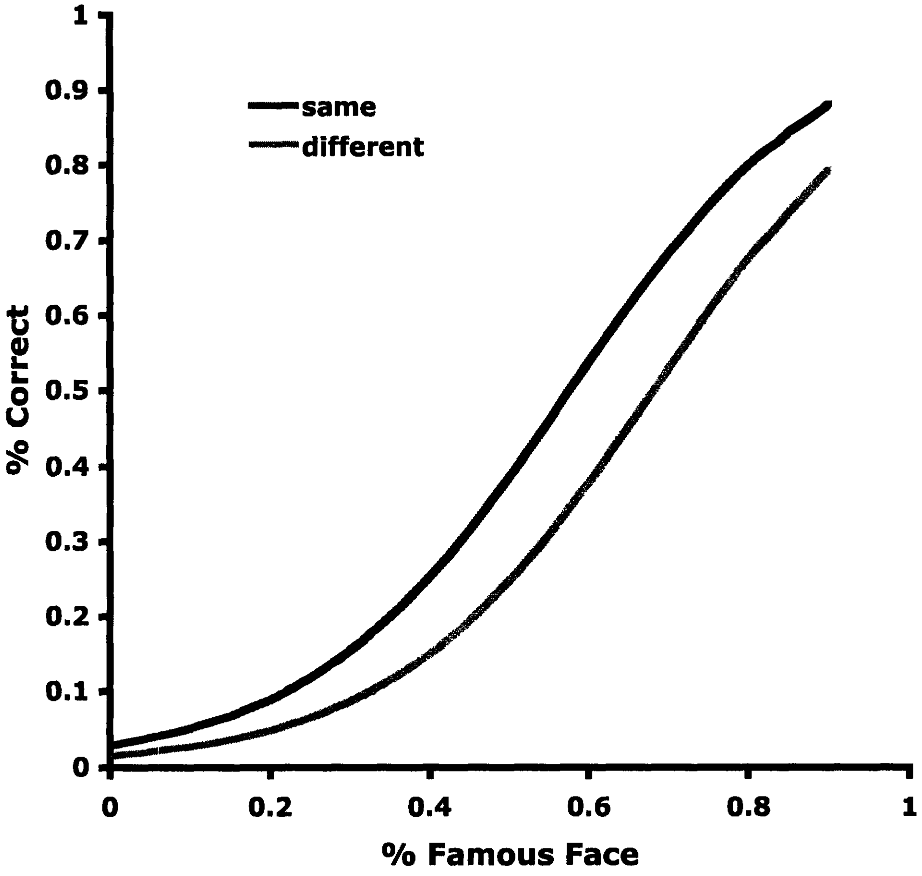


Figure 8: Fitted logistic functions from item analysis of experiment 2. As in experiment 1, subjects needed less of the famous face present in order to recognize it when it was morphed with the adapting (same) stimulus.

As in experiment 1, data were also grouped by item and analyzed. Logistic functions were fit to each item and a parameter representing the shift between the same and different curves obtained for each item. Testing the collection of parameters representing the horizontal shift against zero showed again that faces were better recognized following adaptation to the unknown face they were morphed with ($t(7) = 2.41$, $p = 0.0468$). Figure 8 shows the curves fit to the aggregate data.

Experiment 3:

Experiments 1 and 2 demonstrate that adaptation to unfamiliar faces can influence face recognition in an objective paradigm. In experiment 3 we manipulated the familiarity of the adapting stimuli. Work on a phenomena known as perceptual interference has shown that subjects' hypotheses about an ambiguous or noisy stimulus can interfere with recognition (Bruner & Potter, 1964). In an experiment, subjects were shown highly blurred pictures of objects and asked to recognize them. During the course of the experiment, the amount of blur was slowly reduced and the threshold for identification measured. The finding was that the more blurred the object was initially, the higher the threshold for recognition. The researchers proposed that subjects initial perceptual hypotheses interfered with subsequent recognition. This contrasts with the results of experiments 1 and 2 (and adaptation effects in general) where exposure to an unknown face presumably increased the perception of differences in highly similar faces. However, since the adapting faces were unknown to the subjects they may not have any strong perceptual hypotheses as to the identity of those faces. In experiment 3 we used highly recognizable famous faces (George Bush and Arnold Schwarzenegger) as the adapting faces to see whether this would eliminate or reverse the adaptation effect.

Methods:

Stimuli were created by morphing grayscale photographs of George Bush and Arnold Schwarzenegger with 8 other famous faces (Ben Affleck, Bill Cosby, Elvis, and Sylvester Stallone morphed with Bush; Bill Clinton, Tom Cruise, Dustin Hoffman, and Christian Slater morphed with Schwarzenegger). The experiment was identical to experiments 1 and 2, except that only subjects who could recognize the adapting photos at the start of the experiment did the experiment. 22 subjects recruited from the Stanford undergraduate population participated in return for 5$ payment or undergraduate course credit. 11 subjects adapted to Bush and 11 subjects adapted to Arnold during the experiment.

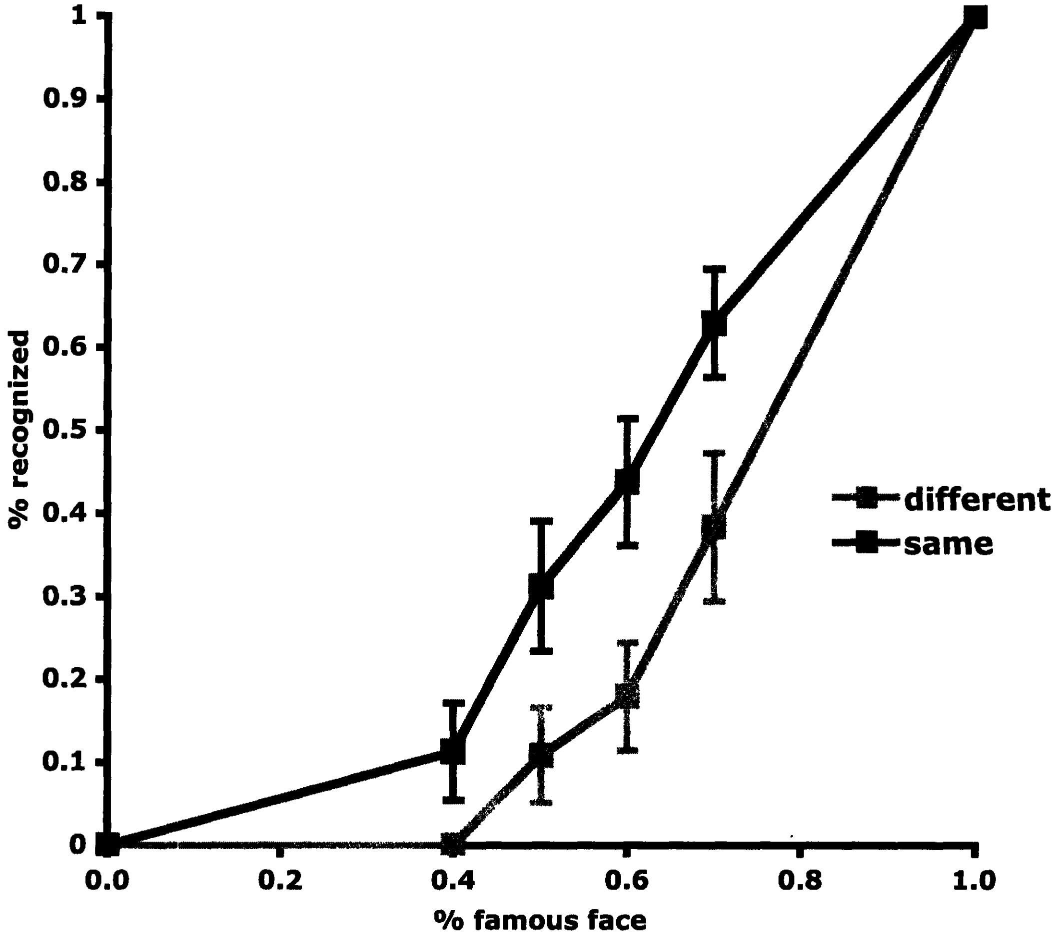


Figure 9: Recognition performance of subjects in experiment 3. Using famous faces as adapting stimuli also produced strong effects on recognition performance at all tested levels of the morphs.

Analysis and Results:

Analysis proceeded exactly as in experiment 1 with faces recognized on the first presentation (40%) prior to adaptation and those faces not recognized at 100% discarded from the analysis. Overall recognition performance in the adapting conditions was lower than experiments 1 and 2, with subjects recognizing 67% of the faces on average (mean number recognized = 5.35, SEM = 0.35).

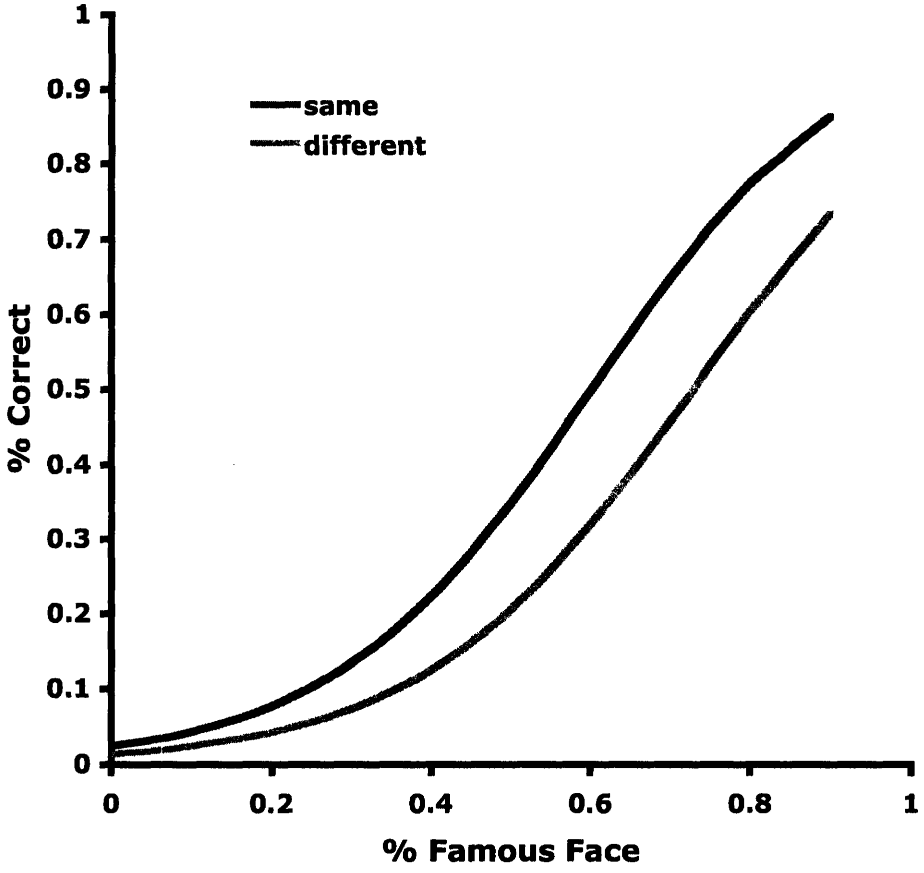


Figure 10: Logistic functions fit to subject data from experiment 3.

Unpaired one-tailed t-tests revealed significant effects of adaptation on all levels of the test faces (40%: t(41)=1.96, $p<0.05$, mean same 11% SE 0.06%. mean different 0% SE 0%; 50%: t(41)=2.3, $p<0.05$, mean same 35% SE 8%, mead different 12% SE 6%; 60%: t(41)=2.22, $p<0.05$, mean same 46% SE 8%, mean different 23% SE 7%; 70%: t(41)=1.89, $p<0.05$, mean same 65% SE 6%, mean different 44%, SE 9%). Note the variance at 40% for the different faces was 0. As a confirmatory analysis, logistic functions were again fitted separately to each subject's same and different data (figures 9 and 10). Estimates of the parameter representing the horizontal shift of the functions were compared using an unpaired t-test and showed a significant effect of adaptation on recognition performance with faces recognized better in the same condition (t(41) =3.2, p<0.005). Logistic functions fit to the item data showed a similar effect, though just barely achieving statistical significance.

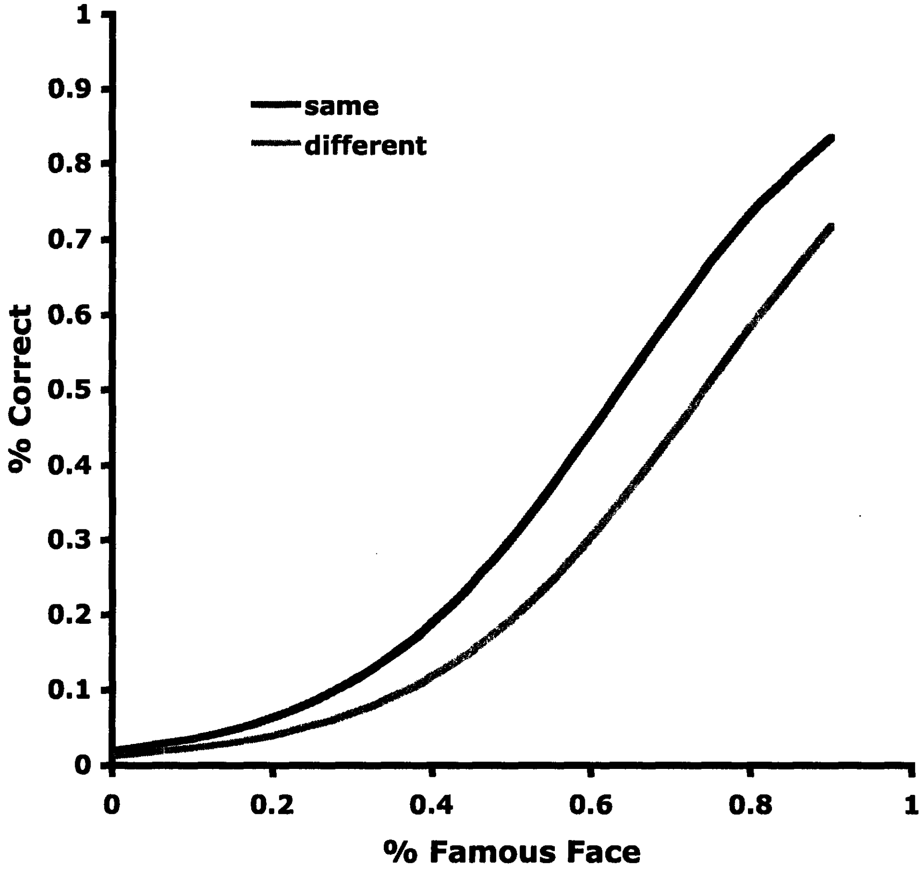


Figure 11: Logistic functions fit to item data from experiment 3. As in experiments 1 and 2, recognition was significantly better in the same condition.

Discussion

The results show that not only does adaptation improve recognition performance by lowering the threshold for recognizing a famous face in a morph, but also that the adaptation affects several trajectories simultaneously. Moreover, experiments 1-3 used

24 different morphs, and the effect of adaptation when considered across these items is highly significant (figure 12).

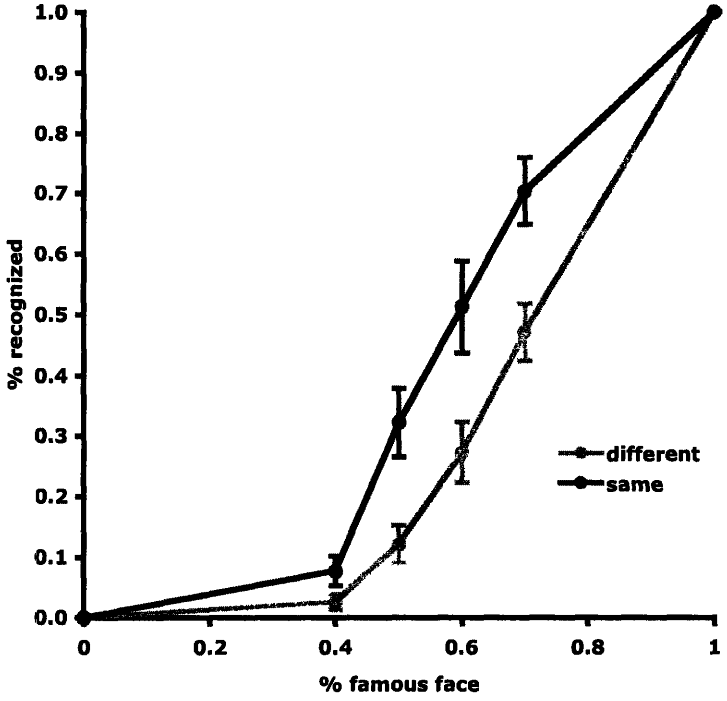


Figure 12: Effect of adaptation on the 24 trajectories used in experiments 1-3. Points are mean recognition across items and error bars are standard error of the mean. Across items, recognition of the famous faces was significantly better when they were morphed with the adapting stimulus (same, purple line) at all levels of the morph tested. This supports the idea that adaptation is effective across many different trajectories.

Paired t-tests comparing performance on each item in the same vs. different adaptation conditions showed superior recognition performance for faces when the adapting stimulus was one endpoint of the morph. This adaptation-driven recognition advantage was found for all test levels (40%: t(23) = 2.13, $p<0.05$, mean same 7.8% SE = 2.4% , mean different 2.6% SE=1.2%; 50%: t(23) = 3.95 $p< 0.001$, mean same 32.3% SE=5.7%, mean different 12% SE = 3.1%; 60%: t(23)=3.95, $p<0.001$, mean same 51.3% SE=7.6%, mean different 27.3% SE = 5%; 70%: t(23)=4.66, $p<0.001$, mean same 70.5%

SE = 5.6%, mean different 47.1% SE=4.7%). The adaptation dependent improvement in recognition supports the idea of identity based aftereffects and provides evidence that these aftereffects are selective and can be obtained along many different unrelated trajectories.

One possible mechanism for the adaptation effect is a shift in the neutral point of a face space as proposed by a number of researchers (Valentine, 1991; Rhodes et al, 2005; Leopold et al, 2001). If the perception on the test faces is accomplished by computing their difference from the norm, then a shift of the norm towards the adapting stimulus would cause faces that lie between the adapting stimulus and another identity on a morph line to be shifted away from the adapting face. This depends on some minimal assumptions about the location of the two faces relative to the norm. For example, if both faces which have been morphed together on the same line originating from the norm (this would correspond to a face and its caricature), than adaptation to the face closer to the center of the space should cause test faces to move closer to the adaptor.

However, strong evidence for a prototype referenced space requires adaptation effects to be selective for the trajectory which passes through the average face, otherwise the results are also consistent with an exemplar model (Rhodes et al, 2006). The results of experiment 3 replicates the findings in 1 and 2 and argues against such selectivity. While it is likely that the average between any two of the faces would be closer to the center of any space in which they were located, it is implausible that all items for which adaptation improved recognition performance were at the center of the space.

Experiments 4 & 5

Numerous studies on face adaptation have shown that aftereffects persist even when differences between the adapting and test stimuli are introduced such as changes in size, viewpoint, and orientation (Jiang et al, 2004; Yamashita et al, 2005; Zhao & Chubb, 2001). This robustness suggests some invariance in the neurons that are being adapted in these paradigms and argues against the possibility that the observed effects are due to some combination of lower-level shape aftereffects (Suzuki & Cavanagh, 1998). Here we test whether or not the aftereffects we measured in experiments 1-3 are obtained when the adapting stimulus is a profile or inverted face. Although face inversion is thought to impair recognition (Yin, 1969), Leopold et al found measurable aftereffects for identity. This contrasts with other results showing that adaptation to gender and distortions are contingent on the up-down orientation of the face (Rhodes et al, 2004; Yamashita et al, 2005).

Methods:

The inversion experiment used the same stimuli as experiments 1 and 2 except that the adapting face was presented upside down. 44 subjects participated in the two versions of this experiment. For the profile experiment, morphs were created between two new unknown faces and 8 famous faces. Arnold Schwarzenegger, Nicholas Cage, Harrison Ford, and Bruce Willis were morphed with one unknown face, and Tom Cruise, Elvis, John Travolta, and Elijah Wood were morphed with the other. The experiment was identical to experiments 1-3 except that subjects adapted to profile views of the unknown faces (see figure 13).

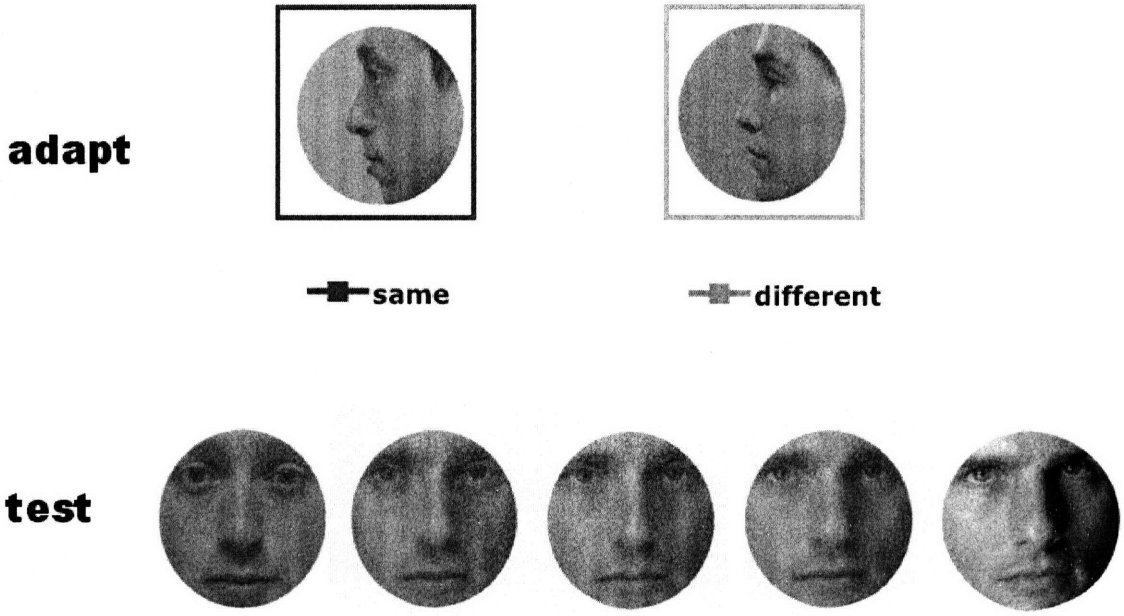


Figure 13. Examples of stimuli used in profile adaptation experiment. the top row shows the two adapting profile. the bottom row shows a the morph between a front view of one unknown face and Tom Cruise.

Analysis and Results:

Adaptation to the inverted faces produced no differences in recognition performance. at any level of morph (maximum t = 0,94, figure 14). Given the absence of any effects the analysis was discontinued at this point.

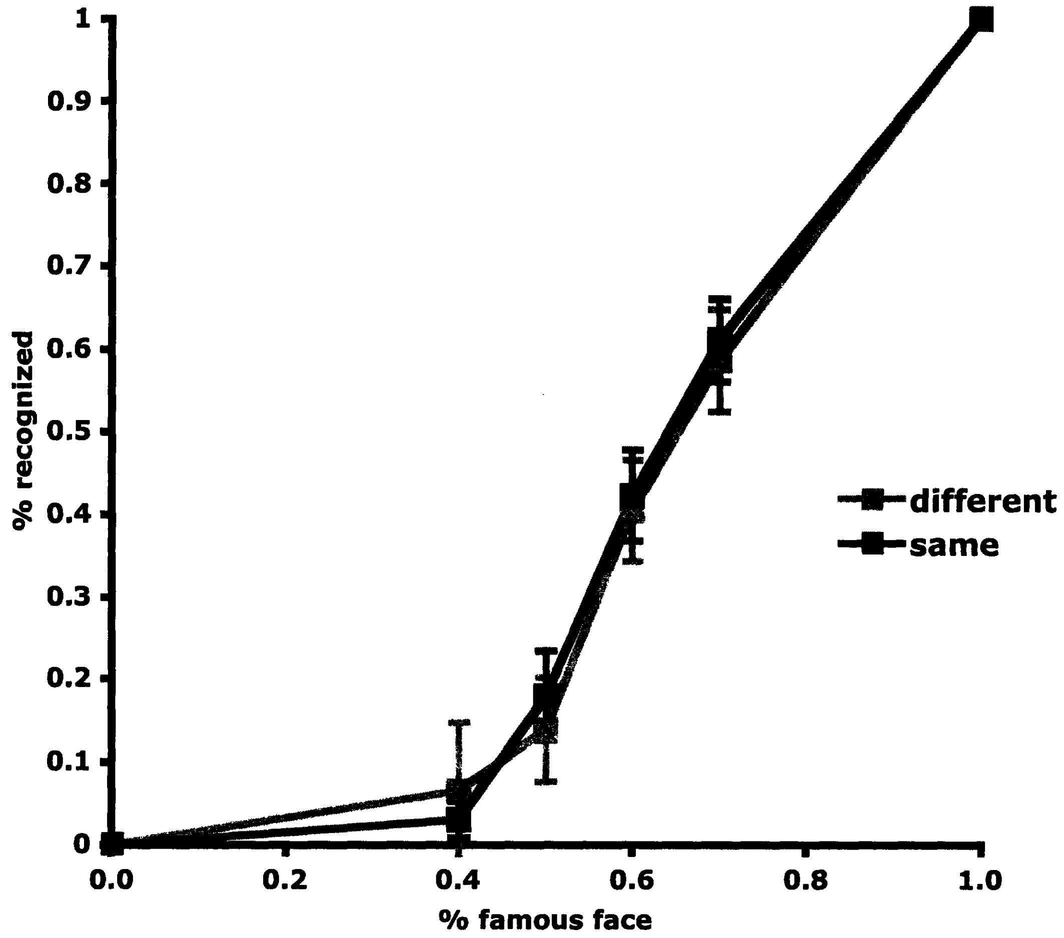


Figure 14: Adapting to inverted unknown faces produced no difference in recognition performance.

Adaptation to the profile views did show a highly significant effect at the 60% famous face (60%: t(37)=2.83, p<0.005, mean same 63% SE=8%, mean different 30% SE=8%; figure 15). No other significant effects were found, with only marginal differences for the 50 and 70% famous faces (50%: t(37)=1.36, p=0.09, mean same 28% SE=7%, mean different 16% SE=5%; 70%: t(37)=1.35, p=0.09, mean same 74%, SE=8%, mean different 58% SE-9%), though these were in the predicted direction.

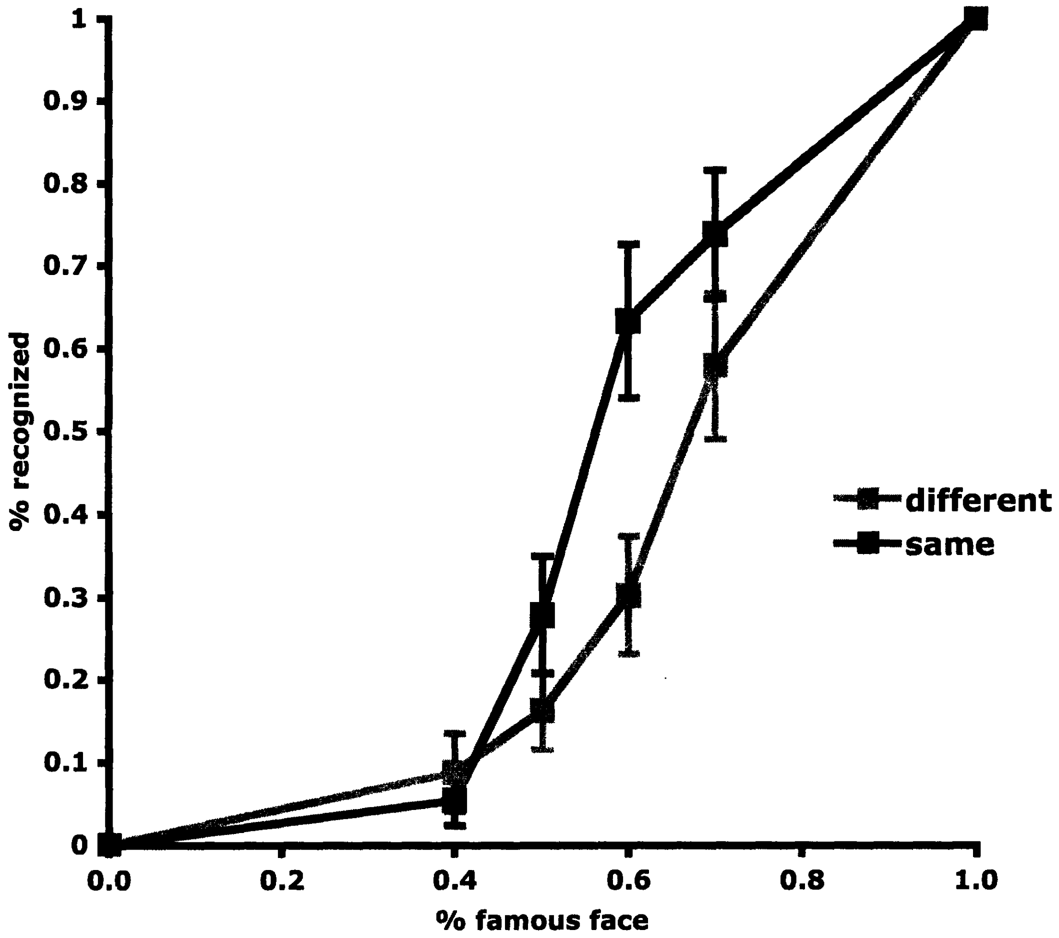


Figure 15: Subjects showed some influence of adaptation to profiles on front view morphs, with same stimuli recognized significantly better at 60% famous face and marginally better at 50 and 70%.

Logistic functions fit to the same and different subject data also revealed only a marginally significant of adaptation on recognition (t(37), p=0.08). An item analysis did not reveal any significant effect.

Discussion:

Adapting to an inverted face did not produce any selective recognition advantage for faces that had been morphed with the upright version of the adapting face. This is consistent with a variety of findings that suggest that adaptation to distortions and gender can be made contingent on the up-down orientation of the adapting stimuli (Rhodes et al, 2004; Witthoft unpublished data). Leopold et al (2001), did find significant identity adaptation to inverted faces. However, the two experiments differ in a number of ways which may account for the divergent findings, including the amount of training (none vs. many hundreds of trials), the stimuli used (photographs vs. reconstructed faces) and task (choose from all the faces you know, vs. a small set of trained examples).

The results from adapting to profile views of the unknown faces are more complex and harder to interpret. While some parts of the analyses provide moderate evidence for transfer between the profile and front view faces, no effect was found in the item analysis. The difference between the profile and front view is a large image transformation, and generally adaptation effects are weakened by changes in the adapting stimulus relative to the test. Though other work conducted in part by the authors has shown that adapting to silhouettes influences gender judgments of front view faces and vice versa (Davidenko et al, 2007), this should be seen as a conservative test. Moreover, there is some evidence that transfer of face adaptation effects across rotations increases when the stimuli are familiar and subjects are able to associate disparate views into a single representation (Jiang et al, 2007). That said it is unlikely that the aftereffects shown in experiments 1-3 are entirely retinotopic, as subjects freely viewed the adapting and test stimuli and numerous other studies have demonstrated that face aftereffects are robust across small image transformations. Future experiments using either smaller image transformations (such as changes in position or 3/4 views) or familiar adapting stimuli will be needed to confirm this position.

General Discussion

Here we have presented results showing that adaptation can selectively enhance the ability to recognize faces. Since subjects did not know which faces out of all the ones they know they would need to recognize and there was a single correct answer for each stimulus the results can not be attributed to response bias. Furthermore, all subjects were tested for recognition on the same set of faces and differed only in the adapting stimulus, demonstrating that any difference between the same and different conditions resulted from the effect of adaptation on extracting information from the test faces. This supports earlier work showing that there are identity specific aftereffects while eliminating concerns that recognition performance may arise from changes in response strategy (Leopold 2001).

Furthermore, faces were chosen arbitrarily, without consideration for where they might be relative to one another in a face space. This leads to two points. First not all of the trajectories can pass through the center of the space (though no doubt the average of any two faces is likely to be closer to the average face than the two faces used to make the average). Second, the advantage for same morphs is numerically larger in the majority of items in experiments 1-3 and produces the predicted shifts in the fitted logistic curves, showing that the adaptation effect is present on several trajectories simultaneously.

Although these results may be consistent with some versions of a prototype referenced encoding scheme which does not require all adaptation effects to pass through the center of the space, they are equally consistent with an exemplar based scheme in which adaptation causes changes such that differences from the adapting stimulus are emphasized. In fact, it is the special directionality of the aftereffect which would allow any face adaptation study to distinguish the exemplar from the prototype view (Tsao & Freiwald, 2006).

In a study closely related to the one found here, Rhodes et al (2006) tested adaptation on trajectories which did not pass through the norm using the same method as Leopold et al (2001). They found initially that, adaptation to these off trajectories produced aftereffects just as large as those found for trajectories passing through the norm, which is consistent with an exemplar view and the results given here. Rhodes et al, then re-measured the baseline following adaptation and found that some learning had taken place (such that subjects were able to identify the target face at a much lower level without adaptation) and that compared to this new baseline the adaptation effects for the off trajectory were now considerably smaller though still significant. No such change in baseline was found for the trajectories which passed through the center of the space, and the authors concluded that faces are indeed represented by a prototype referenced code.

Without disputing the paper in full, this conclusion seems unwarranted. Changes in the baseline for some trajectories and not others could likely be a result of their stimuli rather than some fact about face representation. In the off trajectory case, each adapting stimulus was connected by a morph line to a target face without ever intersecting any other morph line. This means that subjects could easily learn the relationship between the adapt and test stimuli for the off-trajectories and that subjects could learn to recognize the presence of the test in lower levels of the morph. This contrasts with the lines that pass through the center of the space and intersect. The center stimulus is actually the composite of many different faces that subjects are trained on (8 seen in their study) and there is no correct answer for this face. Moreover there is no way to guess what two endpoint faces are connected in the opposite case, since each face is really connected with the center which serves as the other endpoint. Another way of saying this is that as the morph goes from an adapting face to the origin, it moves closer to a face that belongs to multiple trained trajectories. Indeed it is this property that Leopold et al (2001) cited to argue against any claims that their results could have been contaminated by response bias. One possible remedy is to redo the experiment but with only one trajectory that passes through the center. This would remove the differences in the arrangement of the morphs but open up the possibility of response bias influencing all the adaptation judgments. A second possibility would be to have off-trajectories which intersected in a place other than the center of the space as a point of comparison.

Finally, our adaptation effects were obtained without any training on the to-be-recognized faces. This eliminates the possibility that the desired structure of the categories is learned during the training (though this would be an interesting and impressive effect in its own right) and applies just to the stimulus set used in the experiment. That is, it is possible that training people extensively to classify the stimuli may induce a prototype representation for those stimuli. This may seem unlikely, but the Leopold et al (2001) and the Rhodes et al (2006) use different sets of faces to generate their face space but obtain the same result (Rhodes et al, 2005). Perhaps even more surprisingly, Leopold & Bondar (2005) obtained the same behavioral adaptation effect using their stimuli with monkeys. Since the prototype is meant to reflect the average of a person's experiences it is most curious that monkeys would share that prototype and is suggestive that they learned the desired structure during training. Finally, in a follow-up study, Leopold et al (2006) found that the average population response in monkey AIT to their face stimuli increased as they moved away from the average face. This study was conducted with two monkeys using a different face space for each monkey and the authors claimed similar effects in both. In this case the author's acknowledge that the result is a likely artifact of their stimulus set.

Conclusion:

In conclusion, we have shown here that adaptation to one identity can selectively improve recognition using photographs of faces in an objective paradigm. Successful recognition of a famous face in a morph required less of the face to be present when the adapting stimulus was the other endpoint of the morph. Adaptation to profiles showed some effect of transfer to front views suggesting that the adaptation takes place at a relatively high level of processing. Using inverted faces as the adapting stimuli did not produce an any effects on recognition. Improvements in recognition were also found for the majority of the morphs and on multiple trajectories originating from each adapting face showing that identity based aftereffects are not entirely selective for the trajectory passing through the average face.

Chapter 5: Conclusions and Future Directions

In this thesis I have presented three lines of research examining the roles of associative learning, categorization, and adaptation in shaping visual perception. Here I summarize each of these results, recent and subsequent related findings, and possible future directions for research.

Learning and Synesthesia

In chapter 2, I provided evidence that the inducer-concurrent pairings in color-grapheme synesthesia can be learned. Moreover, the data showing that lightness constancy mechanism can influence AED's color matching suggests profound and permanent changes in perception. As noted in the chapter, it is important to say that the results do not address the question of why a particular person becomes a synesthete, and there is some evidence that the condition may be genetic (Baron-Cohen et al, 1996). Rather, our proposal is that for those who are synesthetes, the right kind of stimulus at the right time can determine the grapheme color relationships.

While this is only a single case study, it is important as none other had been described up to this point, and the failure to find a single case of where an environmental stimulus had determined the synesthesia (despite 120 years of research) had led many to believe that learning had no role to play ((Marks, 1975; Ramachandran and Hubbard 2003). Nonetheless, numerous questions remain, two of which I address here.

The first question is can we find some way to more precisely determine the level of the visual system at which the synesthetic colors arise? The answer is possibly, but it requires some discussion of three additional findings from studies of AED.

One interesting feature of her synesthesia is that if a letter is presented in a particular highly saturated color, then she reports that her synesthesia is eliminated. To investigate this we used the matching task described in chapter 2 and asked her to adjust the color of letters until her synesthesia disappeared. The results from some of the matches can be seen in figure 1.

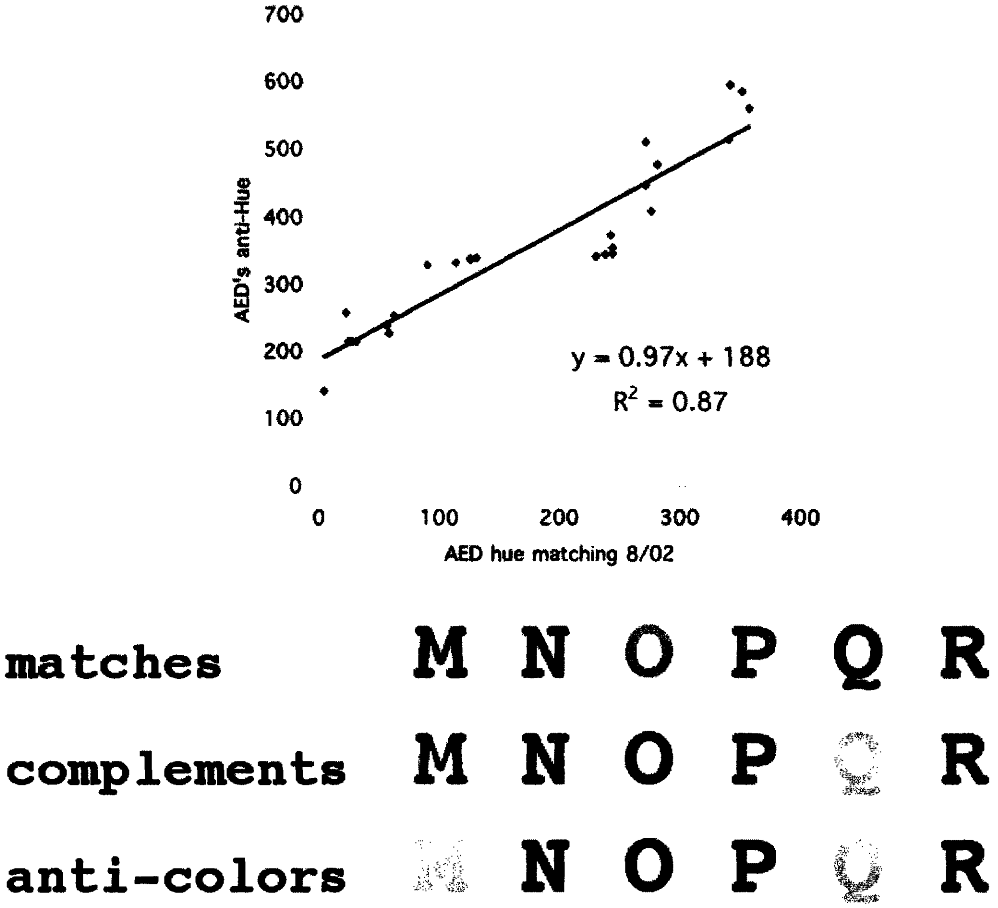


Figure 1: Top shows correlation between hue of letter matches generated by AED which eliminate her synesthesia (y-axis) and matches to synesthetic colors. The two types of matches are highly correlated and with an approximately 188 degree shift around the hue circle. Bottom shows examples of AED matches to synesthetic colors (matches) colors which when mixed with the matches form grey (complements) and AED matches which eliminate her synesthesia (anti-colors).

One striking feature of these matches (what I'll refer to as anti-colors) is that they are roughly complementary to her synesthetic matches, and that when we correlated her anti-colors with her synesthestic colors we find that the hue is highly correlated, but shifted approximately 180 degrees around the hue circle. While we have not carefully characterized whether or not the anti-colors are truly opponent, the data is highly suggestive, and may be a further indication that AED's synesthesia relies on mechanisms used for ordinary color processing.

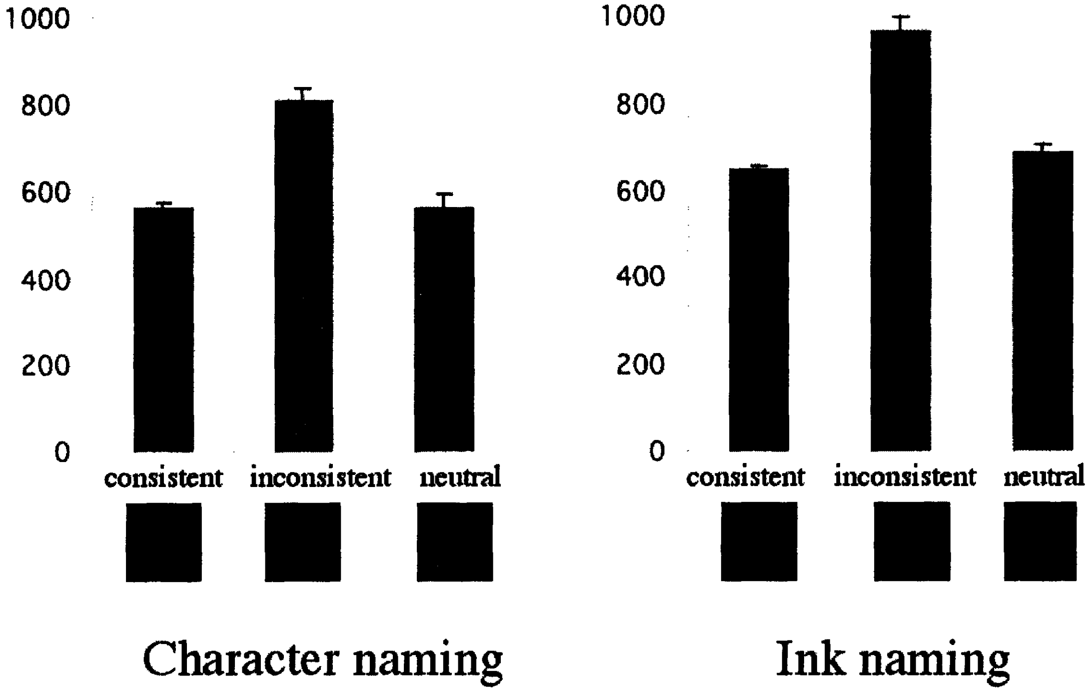


Figure 2: Performance of synesthete AED on Stroop task. Interference as indicated by longer reaction times was apparent when she was asked to name either the letter (left panel) or color of the letter (right panel) presented in an ink inconsistent with her synesthesia. Control subjects tested using the same stimuli showed no effects.

Moreover, we have been able to show that abolishing AED's synesthesia has behavioral consequences. One feature that is common across all types of synesthesia and is in fact taken to be definitional is that the synesthesia is automatic. That is, the concurrent always accompanies the inducer and arises without effort. We addressed the question of automaticity using a modified version of the Stroop paradigm. In one task, AED was presented with letters and asked to name them as quickly as possible. The letters were either colored or gray, with the colors either the same (consistent) as her synesthetic colors or different (inconsistent) from them. In a second experiment she was asked to name the ink color of presented letters and symbols, and again the color of the letters could be either consistent or inconsistent with her synesthesia. As expected, when the stimulus to be named was inconsistent with her synesthesia reaction times increased consistent with response conflict arising from competing cues (namely her synesthestic color and the actual color, figure 2).

Moreover, we were able to divide the inconsistent trials into those where the color was merely different from the synesthetic color and those where it was an anti-color. Another way of putting this is to say that we divided inconsistent trials into that produced synesthesia and those that did not. Interestingly, the absence of the synesthetic color resulted in reduced interference, suggesting that the stroop effect we observed had both a memory (she knows what color the letter should be) and perceptual components (figure 3). Similar effects with other synesthetes using opponent colors in a stroop task have recently been found by another group (Nikolic et al, 2007).

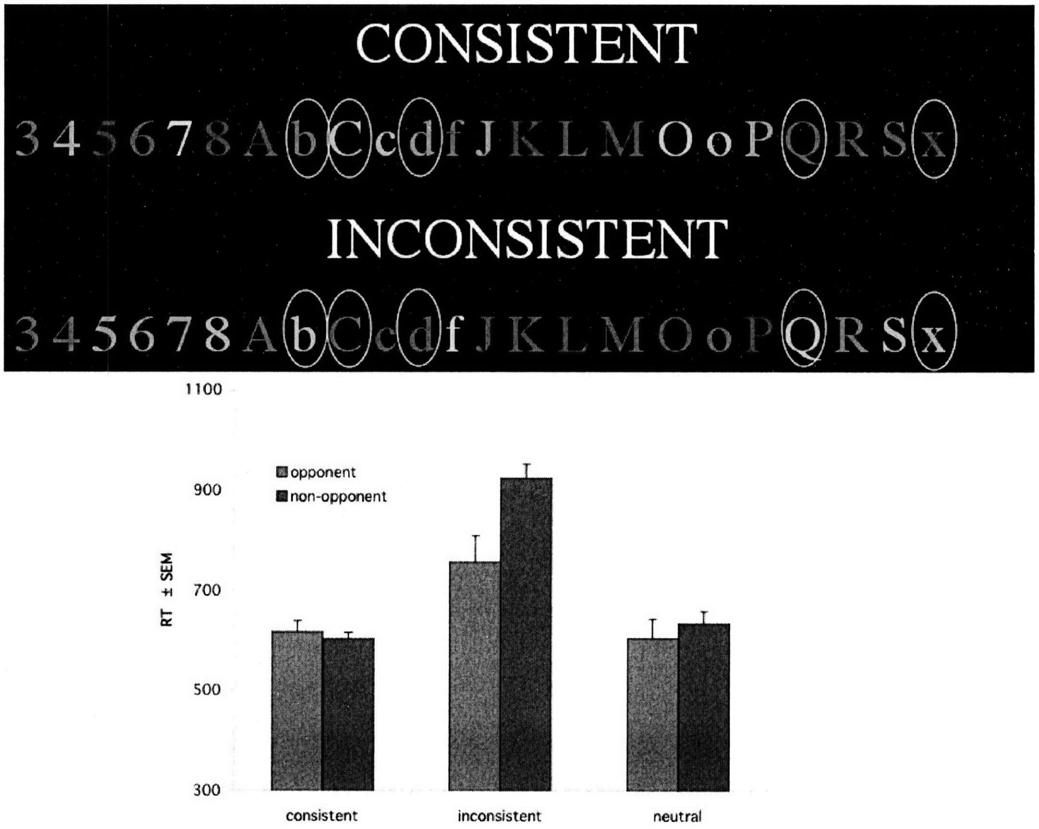


Figure 3: Top panel shows stimuli used in the Stroop experiment with AED. By chance, some of the inconsistent stimuli were also roughly opponent (white circles). Bottom panel shows results from reanalyzing the data splitting between just inconsistent letters and roughly opponent letters. Letters which are both inconsistent and opponent show a reduced amount of interference relative to those which are just inconsistent, suggesting the Stroop effect may have both perceptual and conceptual components. Data are collapsed across ink and letter naming tasks.

While these results provide further evidence that visual mechanisms involved in color perception support AED's synesthesia, they also provide an opportunity to use fMRI to try to directly visualize some of the neural regions involved. We have recently conducted an fMRI experiment in which we showed AED triplets of letters that were either colored or achromatic. In each block the colored letters could either be the same as her synesthesia, different from it, or anti-colored. Her task was to rate the vividness of her synesthesia in response to each triplet. This design allows us to look for regions in the brain where the BOLD response covaries with her perception (figure 4). We also used standard functional localizers which allow us to isolate color responsive regions and to delineate the retinotopic areas of early visual cortex. While we are still in the process of analyzing the data, our expectation is that we will find that some color selective regions will respond in a way correlated with her subjective experience and that the retinotopic analysis will give us a more precise idea where in the brain the synesthetic colors arise (at least in part).

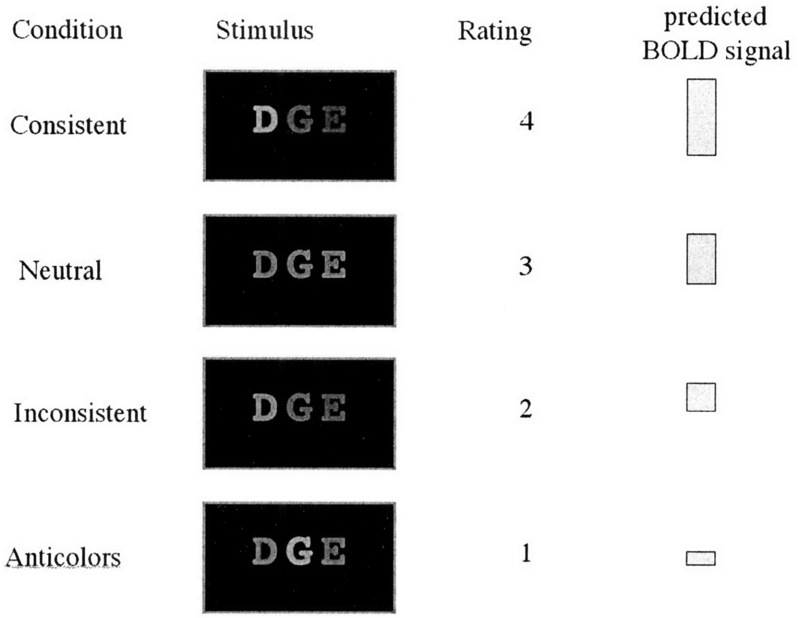


Figure 4: Design of fMRI experiment. Triplets of letters were presented blocked by how they related to the vividness of AED's perceived synesthesia.

A second question to be addressed is whether our case study represents an isolated example, or whether the role of learning in synesthesia is more widespread. In general there has been a recent shift in synesthesia research addressing this question. For example, a finding similar to ours was published simultaneously on a pair of twins who had both learned their synesthetic colors from a childhood toy (Hancock, 2006). Other work has suggested that letter frequency may influence grapheme color matches (Beeli et al, 2007). Moreover, since identifying the initial case, we have come into contact with 5 additional synesthetes who appear to have learned the letter-color pairings from similar sets of refrigerator magnets (figures 5 and 6), including one who has been extensively studied by another group (Kim et al, 2006). The discovery of more examples of this phenomena obviously suggest that the learned pairings may be more widespread than we imagine and further provides an opportunity both for replication of our results and to more generally characterize the underlying mechanisms.

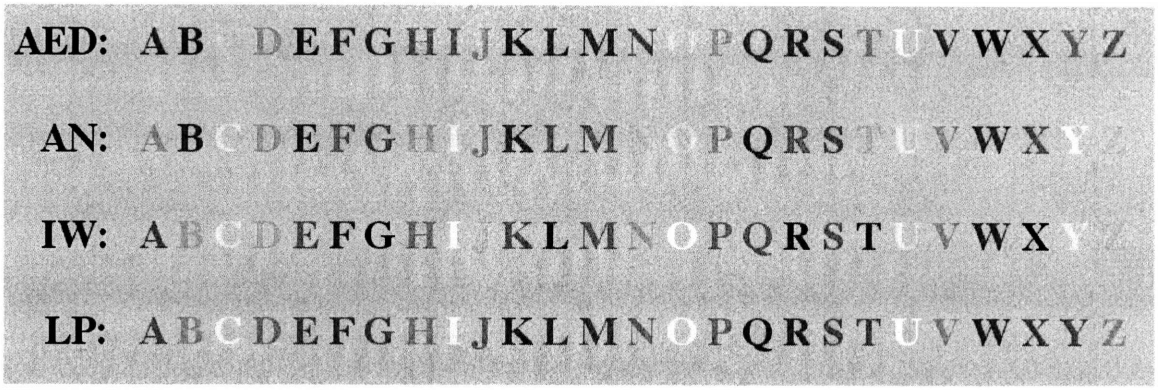


Figure 5: Upper case color matches for 4 synesthetes. Top row is matches for our original subject described in chapter 1. Bottom 3 rows are matches for 3 new synesthetes who also claim to have magnet sets. Matches were obtained using www.synesthete.org and all 3 new subjects scored as synesthetes on both the matching and speeded classification tasks (Eagleman et al, 2007).

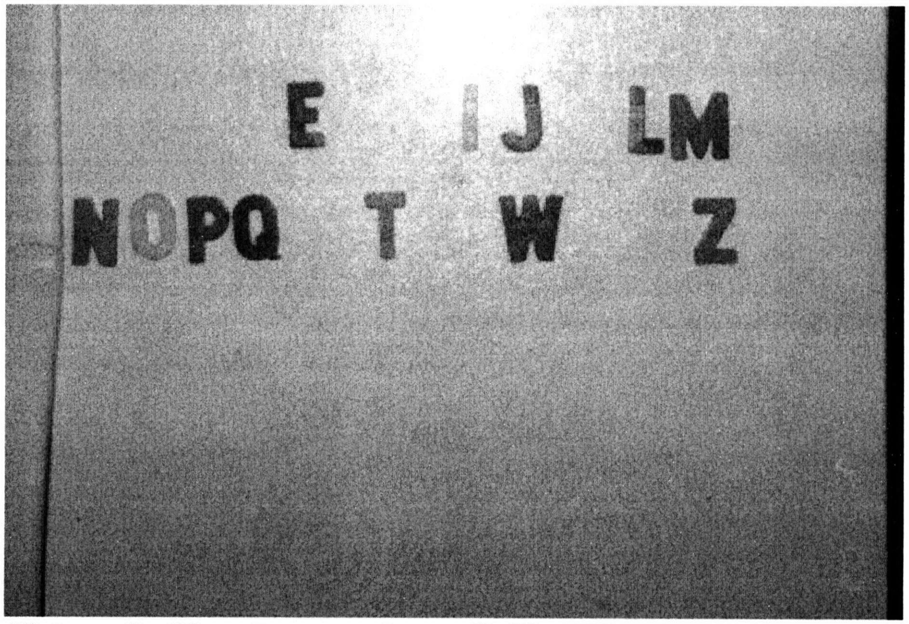


Figure 6: Picture of refrigerator magnets supplied by IW.

Language and Color Perception

In chapter 3, evidence was presented that language can play an online role even in simple color judgments where all the colors are simultaneously available. This effect was found for stimuli that were spatially separated, but not for those that were adjacent suggesting that language may be used to aid the comparison process. The basic finding that linguistic interference can selectively effect color judgments has now been replicated several times both by ourselves in another language (Winawer et al, 2007) and by another group, who found that these effects might only be obtained in the right visual field (Gilbert et al, 2006).

One point worth noting about our results is that they do not address whether or not linguistic categorization results in categorical color perception. To do so, it is necessary to have some means for equating stimulus differences (i.e. the similarity of the color chips) and then showing subsequently that language does or does not distort these distances. One way to get around this is to show that perceived differences change as a function of task as in our experiment or that of Kay & Kempton (1984). That is, showing that performance on the same stimuli differs as a function of experimental condition means that it is not necessary to equate stimulus differences within a condition. This seems a particularly useful approach as it is not clear what the proper method for equating color distances should be.

One possible direction for future research is to investigate whether or not language plays an online role in making judgments about other kinds of stimuli (Roberson & Davidoff, 2000). One idea would be to use morphs between famous faces. Using the same paradigm as in experiment 2 (chapter 3), subjects would be asked to say which of two faces in the bottom of a triad is the same as the face on top. Using morphs would allow us to have both within (same identity) and between category (different identity) lures. Subjects would perform the task alone or while doing a secondary task consisting of spatial or verbal interference. If language plays an online role in categorical judgments more generally, we would expect to find that verbal interference selectively impairs between category trials in this task as well.

Face Adaptation and Recognition

In chapter 4, I introduced a novel objective paradigm for measuring identity based face aftereffects. The results of the experiments showed that identity aftereffects are selective enough to aid recognition of famous faces in face mixtures that also contain the adapting face. This effect was obtained robustly across many items and morph trajectories suggesting that claims that identity based aftereffects pass only through the average face may be premature (Leopold et al, 2001; Rhodes et al, 2006). Our results however, do not allow us to discriminate between exemplar and prototype-based face space hypotheses, and could be seen as consistent with either model (Tsao & Freiwald, 2006). I have also shown evidence of transfer of adaptation from profile faces to front views, suggesting that neurons which specifically code for face adaptation are the source of the effect, and that they are likely fairly high in the visual hierarchy.

The paradigm employed here is particularly useful as it allows us to say that adaptation is making subjects more sensitive to faces that differ from the adapting stimulus (albeit in a particular way). Moreover, it presents what could be a useful method for assessing the degree to which thinking about or imagining relies on perceptual processing mechanisms (Brooks, 1968). Experiments designed to test this kind of question are often undercut by the possibility that subject behavior which looks like a perceptual process is really a strategic decision rather than the result of perceptual processes (Pylyshyn, 2002). However, if subjects were able to improve recognition performance in our paradigm using an imagined face, that would strongly argue that imagining the face adapted the same neurons as perceiving it. Since subjects are unaware of the response alternatives (namely the faces they are trying to recognize) any positive effect must come from a change in how face morphs are processed.

Looking forward, I would like to think this data support the idea that face adaptation may serve a functional role in face perception (Webster et al, 2004; Suzuki, 2005). Faces comprise a highly self similar set of stimuli, and by exaggerating the differences between them, face adaptation may serve to dynamically tune the visual system to be sensitive to the set of features most useful for discriminating and identifying faces.

Conclusions

In this thesis I have presented work designed to assess the different ways in which the visual system and visual representations may change in response to experience. The first set of experiments showed that learning can influence letter-synesthesia pairings. Moreover, this associative learning produced permanent changes in the synesthete's visual experience, suggesting that for some people, the presence of the right stimulus at a critical time can have profound consequences. The second line of experiments demonstrated that language can play an online role even in a relatively simple color discrimination task. In this case the evidence does not show change in specifically visual processing, but that rather linguistic categories (i.e. labels) may be routinely deployed to inform perceptual decisions. The final set of experiments used adaptation to look at the role of recent experience in rapidly adjusting our sensitivity to face information. The recognition paradigm showed that adaptation to identities improved recognition of famous faces that had been 'mixed' with the adapting stimulus, suggesting that adaptation may play a functional role in sensitizing the visual system to features useful for discriminating faces.

Vision does not merely record the world, but rather interprets the patterns of light falling on the retina, transforming these ever changing patterns into the stable constants of our visual experience. Each of the three topics addressed in this thesis, categorization, associative learning, and adaptation, can be thought of as different mechanisms by which new kinds of visual representations can be formed. Although these seem very different, operating at time-scales ranging from many years in the case of categorization, to the moment to moment fine tuning of adaptation, and may be available throughout life or only during early development, the combined effect of these processes is to use experience to organize perception in a way that serves the goals and needs of the perceiver.

REFERENCES

Adelson, E.H. (2000) Lightness perception and lightness illusions. In Gazzaniga, M. (Ed).The New Cognitive Neurosciences. Cambridge, MA. MIT Press, Ch. 24, 339-351.

Ahissar, M., & Hochstein, S. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. Neuron, 36(5), 791-804.

Barlow, H.B. (1990) A theory about the function role and synaptic mechanism of visual aftereffects. In C. Blakemore (Ed.) Vision: Coding and Efficiency. Cambridge, Cambridge University Press.

Baron-Cohen, S., Wyke, M.A., & Binnie, C. (1987) Hearing words and seeing colours: and experimental investigation of a case of synesthesia. Perception,16: 761-767.

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., & Bolton, P. (1996) Synaesthesia: prevalence and familiarity. Perception, 25: 1073-1079.

Beeli, G., Esslen, M., Jancke, L. (2007) Frequency correlates in grapheme-color synesthesia. Psychological Science 18(9), 788-792.

Berlin, B., & Kay, P. (1999). Basic Color Terms: Their Universality and Evolution. The David Hume Series of Philosophy and Cognitive Science Reissues. Port Chester, NY: Cambridge University Press.

Blanz, V., & Vetter, T. (1999) A morphable model for the synthesis of 3d faces. SIGGRAPH Conference Proceedings, 187-194.

Bornstein, M. H., Kessen, W., & Weiskopf, S. (1976). Color vision and hue categorization in young human infants. Journal of Experimental Psychology: Human Perception & Performance, *2*(1), 115-129.

Bornstein, M. H., & Monroe, M. D. (1980). Chromatic Information-Processing - Rate Depends on Stimulus Location in the Category and Psychological Complexity. Psychological Research-Psychologische Forschung, 42(3), 213-225.

Brooks, L.R. (1968) Spatial and verbal components of the act of recall. Canadian Journal of Psychology 22(5), 348-68.

Bruner, J.S., & Potter, M. (1964) Interference in visual recognition. Science 144, 424-5.

Cabeza, R., Bruce, V., Kato, T., & Oda, M. (1999) The prototype effect in face recognition: Extension and Limits. Memory and Cognition, 27, 139-151.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). Psyscope: A new graphic interactive environment for designing psychology experiments. Behavioral Research Methods, Instruments, and Computers, 25(2), 257-271.

Cox, D., Meier, P., Oertelt, N., & DiCarlo, J. (2005) Breaking position invariant object recognition. Nature Neuroscience 8(9), 1145-47.

Cutzu, F. & Edelman, S. (1996). Faithful representation of similarities among three-dimensional shapes in human vision. Proceedings of the National Academy of Sciences, 93, 12046-12050.

Davidenko, N., Witthoft, N., & Winawer, J. (2007) Gender aftereffects in face silhouettes reveal face-specific mechanisms. OPAM conference paper to appear in Vision Research.

Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories in a stone-age tribe. Nature, 398(6724), 203-204.

Davies, I., Sowden, P. T., Jerret, D. T., Jerret, T., & Corbett, G. G. (1998). A cross-cultural study of English and Setswana speakers on a colour triads task: A test of the Sapir-Whorf hypothesis. British Journal of Psychology, 89, 1-15.

Fang, F., & He, S. (2005). Viewer-centered object representation in the human visual system revealed by viewpoint aftereffects. Neuron, 45(5), 793-800.

Franklin, A., Clifford, A., Williamson, E., & Davies, I. (2005). Color term knowledge does not affect categorical perception of color in toddlers. Journal of Experimental Child Psychology, 90(2), 114-141.

Franklin, A., Pilling, M., & Davies, I. (2005). The nature of infant color categorization: evidence from eye movements on a target detection task. Journal of Experimental Child Psychology, 91(3), 227-248.

Galton, F. (1880) Visualised Numerals. Nature, 22, 494-495.

Galton, F. (1907) Inquiry into Human Faculties and its Development. New York . E.P. Dutton.

Gibson, J.J., & Radner, M. (1937) Adaptation, after-effect, and contrast in the perception of tilted lines. I. Quantitative studies. Journal of Experimental Psychology 20, 453-67.

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. Proceedings of the National Academy of Sciences of the United States of America, 103(2), 489-494.

Goldstone, R. (1994). Influences of categorization on perceptual discrimination. Journal of Experimental Psychology: General, 123(2), 178-200.

Grossenbacher, P., & Lovelace, C. (2001) Mechanisms of synesthesia: cognitive and physiological constraints. Trends in Cognitive Sciences, 5: 36-41.

Hancock, P. (2006) Monozygotic twins' colour-number associations: A case study. Cortex 42, 147-150.

Haijiang, Q., Saunders, J.A., Stone, R.W., & Backus, B.T. (2006) Demonstration of cue recruitment: Change in visual appearance by means of Pavlovian conditioning. Proceedings of the National Academy of Sciences 103(2), 483-8.

Held, R. (1980) The rediscovery of adaptability in the visual system: Effects of extrinsic and intrinsic chromatic dispersion. In Harris, C.S. (Ed) Visual Coding and Adaptability. Hillsdale, New Jersey: Lawrence Erlbaum Associates. 69-94.

Jiang, F. Blanz, V. & O'Toole (in press, 2007). The role of familiarity in three-dimensional view transferability of face identity adaptation. Vision Research, 47(4), 525-531

Karni, A. (1996) The acquisition of perceptual and motor skills: a memory system in the adult human cortex. Cognitive Brain Research 5(1), 39-48.

Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? American Anthropologist, 86, 65-78.

Kay, P., & Regier, T. (2003). Resolving the question of color naming universals. Proceedings of the National Academy of Sciences, U S A, 100(15), 9085-9089.

Kim, C., Blake, R., & Palmeri, T.J. (2006) Perceptual interaction between real and synesthetic colors. Cortex, 42, 195-203.

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. Nature Neuroscience, 4(1), 89-94.

Leopold, D. A. & I. Bondar. (2005) Adaptation to complex visual patterns in humans and monkeys. in Clifford, C. W. G., & Rhodes, G. (Eds). Fitting the mind to the world : adaptation and after-effects in high-level vision (1st ed.). Oxford ; New York: Oxford University Press. 189-212.

Leopold, D.A., Bondar, I., & Giese, M.A. (2006) Norm-based face encoding by single neurons in monkey inferotemporal cortex. Nature, 442(7102), 572-5.

Lewis, M.B. (2004) Face-Space-R: Towards a unified account of face recognition. Visual Cognition 11(1), 29-69.

Logothetis, N.K., Pauls, J., & Poggio, T. (1995) Shape representation in the inferior temporal cortex of monkeys. Current Biology 5(5), 552-563.

Macevoy, S.P., & Paradiso, M.A. (2001) Proceedings National Academy of Sciences, 98: 8827-8831.

Marks, L.E. (1975) On Colored-hearing Synesthesia: Cross-Modal Translations of Sensory Dimensions. Psychological Bulletin, 82: 303-331.

Maurer, D. (1997) Neonatal synesthesia: implications for the processing of speech and faces. In Baron-Cohen S and Harrison J (Eds) Synaesthesia: Classic and Contemporary Readings. Cambridge, MA:, Blackwell Publishers,, Ch. 12, 224-242.

Mills, C., Viguers, M., Edelson, S., Thomas, A., Simon-Dack, S., & Innis, J. (2002) The color of two alphabets for a multilingual synesthete. Perception, 13: 1371-1394.

Newhall, S. M., Burnham, R. W., & Clark, J. R. (1957) Comparison of successive with simultaneous color matching. Journal of Optical Society of America, 47(1), 43-56.

Nikolic, D., Lichti, P., & Singer, W. (2007) Color-opponency in synesthetic experiences. Psychological Science, 18(6):481-486.

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

Ozgen, E., & Davies, I. R. (2002). Acquisition of categorical color perception: a perceptual learning approach to the linguistic relativity hypothesis. Journal of Experimental Psychology: General, 131(4), 477-493.

Palmer, J., Huk, A., & Shadlen, M. (2005) The effect of stimulus strength on the speed and accuracy of a perceptual decision. Journal of Vision 5(5), 376-404.

Palmeri, T.J., Blake, R., Marois, R., Flanery, M., & Whetsell, W. (2002) The perceptual reality of synesthetic colors. Proceedings of the National Academy of Sciences, 99: 4127-4131.

Pilling, M., & Davies, I. R. (2004). Linguistic relativism and colour cognition. British Journal of Psychology, 95(Pt 4), 429-455.

Pilling, M., Wiggett, A., Ozgen, E., & Davies, I. R. (2003). Is color "categorical perception" really perceptual? Memory and Cognition, 31(4), 538-551.

Pinker, S. (1994). The Language Instinct (1rst ed.). New York, NY: W. Morrow & Co.

Posner, M.I., & Keele, S.W. (1968) On the genesis of abstract ideas. Journal of Experimental Psychology, 77(3), 353-63.

Pylyshyn, Z.W. (2002) Mental imagery: In search of a theory. Behavioral and Brain Sciences 25, 157-238.

Ramachandran, V.S. and Hubbard, E.M. (2001a) Synesthesia-A window into perception, thought and language. Journal of Consciousness Studies, 8: 3-34.

Ramachandran, V.S., & Hubbard, E.M. (2001b) Psychophysical investigations into the neural basis of synaesthesia. Proceedings of the Royal Society of London, Series B, 268: 979-983.

Ramachandran, V.S., & Hubbard, E.M. (2003a) Hearing Colors, Tasting Shapes. Scientific American, 52-59.

Ramachandran, V.S., & Hubbard, E.M. (2003b) The phenomenology of synaesthesia. Journal of Consciousness Studies, 10: 49-57.

Regier, T., Kay, P., & Cook, R. S. (2005). Focal colors are universal after all. Proceedings of the National Academy of Sciences, U S A, 102(23), 8386-8391.

Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W., & Nakayama, K. (2003). Fitting the mind to the world: face adaptation and attractiveness aftereffects. Psychological Science, 14(6), 558-566.

Rhodes, G., Jeffery, L., Watson, T. L., Jaquet, E., Winkler, C., & Clifford, C. W. (2004). Orientation-contingent face aftereffects and implications for face-coding mechanisms. Current Biology, 14(23), 2119-2123.

Rhodes, G., Robbins, R., Jacquet, E., McKone, E., Jeffrey, L., & Clifford, CWG. (2005). in Clifford, C. W. G., & Rhodes, G. (Eds). Fitting the mind to the world : adaptation and after-effects in high-level vision (1st ed.). Oxford; New York: Oxford University Press. 213-240.

Rhodes, G., & Jeffrey, L. (2006) Adaptive norm-based coding of facial identity. Vision Research 46, 2977-2987.

Rich, A., Mattingley, J.B. (2002). Anomalous perception in synesthesia: A cognitive neuroscience perspective. Nature Reviews Neuroscience, 3, 43-52.

Roberson, D., Davidoff, J., & Braisby, N. (1999). Similarity and categorisation: neuropsychological evidence for a dissociation in explicit categorisation tasks. Cognition, 71(1), 1-42.

Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: the effect of verbal interference. Memory and Cognition, 28(6), 977-986.

Roberson, D., Davies, I., & Davidoff, J. (2000). Color categories are not universal: replications and new evidence from a stone-age culture. Journal of Experimental Psychology General, 129(3), 369-398.

Roberson, D., Davidoff, J., Davies, I. R., & Shapiro, L. R. (2004). The development of color categories in two languages: a longitudinal study. Journal of Experimental Psychology General, 133(4), 554-571.

Roberson, D., Davidoff, J., Davies, I. R., & Shapiro, L. R. (2005). Color categories: evidence for the cultural relativity hypothesis. Cognitive Psychology, 50(4), 378-411.

Sinha, P., & Poggio, T. (1996) Role of learning in three-dimensional shape perception. Nature, 364(5), 460-3.

Smilek, D., Dixon, M.J., Cudahy, C., & Merikle, P.M. (2001) Synesthetic photisms influence visual perception. Journal of Cognitive Neuroscience, 13: 930-936.

Suzuki, S. & Cavanagh, P. (1998) A shape-contrast effect for briefly presented stimuli. Journal of Experimental Psychology: Human Perception and Performance. 24(3), 1315-1341.

Suzuki, S. (2005) High-level pattern coding revealed by brief shape aftereffects. in Clifford, C. W. G., & Rhodes, G. (Eds). Fitting the mind to the world : adaptation and after-effects in high-level vision (1st ed.). New York: Oxford University Press. 135-72

Troje, N.F., Sadr, J., Henning, G., & Nakayama, K. (2006) Adaptation aftereffects in the perception of gender from biological motion. Journal of Vision 6, 850-7.

Tsao, D., Freiwald, W.A. (2006) What's so special about the average face? Trends in Cognitive Sciences. 10(9), 391-3.

Valentine, T. (1991) A unified account of the effects of distinctiveness, inversion, and race on face recognition. Quarterly Journal of Experimental Psychology 43, 161-204.

Valentine, T. (2001) Face-space models of face recognition. In MJ Wenger & JT Townsend (Eds.), Computation, geometric, and process perspectives of faction cognition: Contexts and Challenges. Hillsdale, New Jersey: Earlbaum.

Wandell, B. (1995) Foundations of Vision. Stamford, CT: Sinauer Associates.

Wallach, H., & Austin, P. (1954) Recognition and Localization of Visual Traces. American Journal of Psychology, 67(2), 338-340.

Wallis, G. & Bulthoff, H. (1999) Learning to recognize objects. Trends in Cognitive Sciences 3(1), 22-31.

Ward, J. & Simner, J. (2003) Lexical-Gustatory synaesthesia: linguistic and conceptual factors. Cognition, 89: 237-261.

Ward, J., Huckstep, B., & Tsakanikos, E. (2006) Sound-colour-synaesthesia: To what extent does it use mechanisms common to us all? Cortex, 42, 264-280.

Ward, J., & Mattingley, J.B. (2006) Synasthesia: An overview of contemporary findings and controversies. Cortex, 42, 129-136.

Webster, M. A., & MacLin, O. H. (1999). Figural aftereffects in the perception of faces. Psychonomic Bulletin & Review, 6(4), 647-653.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. Nature, 428(6982), 557-561.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007) The Russian Blues: Effects of language on color discrimination. Proceedings of the National Academy of Sciences, 104(19), 7780-5.

Witthoft, N., Winawer, J., Wu, L., Frank, M., Wade, A., Boroditsky, L. (2003). Effects of Language on Color Discriminability. Paper presented at the Cognitive Science Society, Boston, Ma.

Yamashita, J. A., Hardy, J. L., De Valois, K. K., & Webster, M. A. (2005). Stimulus selectivity of figural aftereffects for faces. Journal of Experimental Psychology: Human Perception & Performance, 31(3), 420-437.

Yin, R.K. (1969) Looking at upside-down faces. Journal of Experimental Psychology, 81(1), 141-5.

Zhao, L., & Chubb, C. (2001). The size-tuning of the face-distortion after-effect. Vision Research, 41(23), 2979-2994.