Learning to see:
The early stages of perceptual organization

by

Yuri Ostrovsky

A.B. Computer Science
Harvard College, 1998

SUBMITTED TO THE DEPARTMENT OF BRAIN AND COGNITIVE SCIENCES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN NEUROSCIENCE
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

SEPTEMBER 2010

Signature of Author: ____
Department of Brain and Cognitive Sciences
September, 2010
Certified by: ____
Pawan Sinha
Associate Professor of Computational Neuroscience
Thesis Supervisor
Accepted by:____
Earl Miller
Picower Professor of Neuroscience
Director, Brain and Cognitive Sciences Graduate Program

Learning to see:
The early stages of perceptual organization

by

Yuri Ostrovsky

Submitted to the Department of Brain and Cognitive Sciences
September, 2010 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Neuroscience

ABSTRACT

One of the great puzzles of vision science is how, over the course of development, the complex visual array comprising many regions of different colors and luminances is transformed into a sophisticated and meaningful constellation of objects. Gestaltists describe some of the rules that seem to govern a mature parsing of the visual scene, but where do these rules come from? Are they innate—endowed by evolution, or do they come somehow from visual experience? The answer to this question is usually confounded in infant studies as the timelines of maturation and experience are inextricably linked. Here, we describe studies with a special population of late-onset vision patients, which suggest a distinction between those capabilities available innately and those which are crafted via learning from the visual environment. We conclude with a hypothesis, based on these findings and other evidence, that early-available common fate motion cues provide a level of perceptual organization which forms the basis for the learning of subsequent cues.

Thesis Supervisor: Pawan Sinha
Title: Associate Professor of Computational Neuroscience

Table of Contents

Chapter 1. Introduction

Vision – the ability to "see"—seems so natural and seamless to us that a casual student of the mind may not even consider this faculty as requiring study. The lay observer often considers the eye as the source of vision, matching its processed images to memories in the brain. Vision is considered the domain of optometrists and ophthalmologists, not of psychologists and neuroscientists. If further thought is given, the role of the brain in this conceptualization is deemed to be that of a "matchmaker" between the output of the eye and the stored visual memories and their semantic tags: Vision is little more than an output of color values and their locations.

The more informed or thoughtful observer realizes that vision is tremendously more complex than mapping points of light to stored memories of these points. One must infer the structure of a 3-dimensional world from a 2D projection, a process far beyond simple "matching." Objects, textures, and faces can be viewed from an infinite number of perspectives under wildly varying conditions of illumination and occlusion. It is the role of the brain to solve this incredibly difficult problem of "parsing" an image into interpretable real-world "stuff".

The Gestaltist tradition in vision science (Wertheimer, 1923) inspired a search for "rules" that might govern this visual parsing. Kanizsa (1979) catalogued a large number of conditions and "cues" (e.g., symmetry, regularity, continuity; and more notably, when these cues fail as well) from this research that seem to serve to disambiguate objects from background and from other objects, but he recognized that even this substantial list was plagued with myriad exceptions and could only explain a subset of the stimuli that normal adults easily decompose into accurate interpretations.

One of the most important conceptual frameworks of the computational problem involved in the parsing of the visual world was given by Marr (1982). This treatise highlights the complexity of visual parsing from the point of view of the algorithms and calculations that such a system must in some form implement. Marr decomposed the process into three stages: the primal sketch, identifying properties of the image such as edges and regions; the 2.5D sketch, interpreting viewer-center aspects of the scene such as surface geometry; and the 3D model, providing a full reconstruction of the objects independent of viewpoint. Although the details and validity of this model are debated, it serves the invaluable purpose of providing at least one concrete perspective on the problems that need to be solved by the brain in order to process an image into a model of the physical world. It is this process that forms the foundation of what we generally think of as “vision”, as opposed to the optical properties of the eye.

Given the complexity of vision, the relative ease of our visual parsing accomplishments, and the observation that even infants seem to have this ability, it is tempting to conclude that evolution has endowed us with this ability, just as it has formed the eye that provides us with optical input. In this view, infants are either born with a genetically pre-programmed ability to interpret visual scenes into 3D scenes, or their brains mature over a short period of time, allowing this ability to emerge in a structured manner. Although it may not be obvious to the parent trying to differentiate between these two views merely by observation--since general intelligence, motor skills, and other abilities are developing simultaneously—infant vision (both in terms of acuity and the ability to parse the visual scene) is not equivalent to its adult counterpart, and we discuss the differences below. This negates the hypothesis that vision is immediately available in its full adult form. However, the “maturation” hypothesis—i.e., that visual skills develop as the brain matures through a genetic program— is a viable one.

The alternative viewpoint to visual development is that experience with the visual and physical world allows an infant to “learn” visual abilities. Although not as intuitive with respect to our personal experiences as the genetically-endowed version of events, there is ample evidence that at least some visual learning occurs later in life solely from experience with the visual environment, including, as mere examples, the increased ability to detect the motion of random dots in a particular direction on which the observer has been trained (Ball & Sekuler, 1987) and the parsing of object contours that were invisible prior to exposure to the ground-truth object concept (e.g., the Dalmatian dog illusion, to be discussed later (Gregory, 1970)) . Thus, there is evidence that the brain is “plastic” or changeable into adulthood even when it comes to visual processing, though we have not yet shown to what extent this plasticity applies to the general “rules” of vision, such as those mentioned earlier.

This question of whether visual perception is learned or innate is the subject of endless debate and has rich historical roots dating back to at least the late 17th century when the distinction between nativism and empiricism was sharply highlighted in the works of philosophers like Locke, Berkeley and Hume.

The answer to the question of the origins of visual processing will likely not be a simple one. Vision encompasses a number of complex processes, and some of these processes may be completely genetically pre-determined (i.e,. innate), others may involve an interaction between innate and experiential factors, and yet others may be almost entirely a result of learning from the environment. It is the goal of this chapter to differentiate among these processes and to summarize the evidence for their provenance; it is the goal of this thesis to provide novel evidence that parsing abilities may be more dependent on learning than is usually presumed and to outline a possible method for bootstrapping this learning.

1.1 "Maturation" versus "development"

Although the words "maturation" and "development" are often used interchangeably, here we must adopt a more precise usage in order to make subtle but important distinctions in our discussion of vision acquisition. Both words, in the context of neuroscience, usually refer to the changes that occur from birth to adulthood, during which time physiology (including neurophysiology) and cognitive abilities are progressing towards an adult state. We refer to this multi-faceted progression generally as "development." We refer to the physiological and anatomical changes that occur specifically in the brain and in the rest of the body as "maturation." Thus, maturation is a subset of development.

To illustrate, let us consider the "development" of walking. An infant must practice the use of its limbs repeatedly before mastering this skill. This occurs through learning and is accelerated through repetition (P. R. Zelazo, N. A. Zelazo, & Kolb, 1972). However, walking is also dependent on maturation – the muscles of the legs and torso must be sufficiently strong to support the infant. Although these muscles also get stronger through use, in these early stages general growth is required first, and this growth of course is genetically preordained. Physical growth (muscle mass and leg size), a function primarily of maturation, plays a role in the enablement of learned processes in walking dynamics (Thelen, Fisher, & Ridley-Johnson, 1984), and thus both mechanisms contribute to general walking development.

Maturation is the product primarily of genetic programming (i.e., intrinsic factors), whereas *development* is the product of *both* maturational processes as well as any changes that occur through practice, learning, or environmental interaction (extrinsic factors). That is not to say that maturation is completely independent of experience. Our muscles generally become stronger as we progress through adolescence due to the maturational process of puberty and general growth, but not if they experience prolonged disuse. A similar pattern may be observed in the brain. The development of ocular dominance columns are a characteristic of visual neuroanatomy whose beginnings are evident immediately after birth, yet visual deprivation in one eye can have permanent deleterious effects on this neuroanatomy (Hubel, Wiesel, & LeVay, 1977; Wiesel & Hubel, 1965b), suggesting some role of experience in a process with a maturational origin.

Another distinction can be provided using a computer metaphor. An electronic device with only read-only memory can be thought of as a pure incarnation of a "genetic" (or "innate") program, where all aspects of its behavior are predetermined. The resulting behavior would in this case be innate as well. Now consider a device that implements a generic learning algorithm and learns a variety of behaviors by extracting information from its sensory environment, the results of which it stores in its working memory. In this example, the learning algorithm is "innate," while the behaviors are clearly acquired through learning. Some algorithms may even be able to modify themselves, thus making the concept of "innate" less intuitive, as the algorithm that ultimately gives rise to the learning of a behavior itself came from a learning process. Finally, consider a system somewhere in between these two, in which an innate algorithm implements a behavior, but the algorithm requires experience with the environment to set parameters or to calibrate some details of its computations. Is the resulting behavior now innate or learned? The answer seems to lie somewhere between the two possibilities.

Thus, we see that this line between maturation and the components of development that are not maturational (usually termed "practice" for muscular development and "learning" for neurological and cognitive development) is not necessarily a clear one. In one extreme sense, every one of our cognitive skills depend on maturation, as they would not exist without a certain pre-wiring that enables the acquisition of these skills; in the other extreme, perhaps none of our skills are purely maturational, as complete sensory deprivation from birth would likely impede their acquisition as well. Perhaps it would be more accurate to discuss a continuum, rather than making a categorical distinction, but for simplicity in our discussion of visual development, we will usually employ the label "maturation" for processes that are largely determined innately and "learning" for processes that are acquired more flexibly and are to a significant extent dependent on experience.

1.2 A brief history of visual development research

Research in the field of "visual development" provides us with a robust foundation of prior work on which to build. Here, we present some key findings on the progression of visual abilities. We begin with a survey of physiological, anatomical and brain imaging studies on brain plasticity. Next, we summarize the behavioral literature on visual development. In order to provide some structure, as well as a useful theoretical distinction, the behavioral findings will be divided into two sections: "lower-level vision" and "higher-level vision." Lower-level vision will refer to aspects of vision such as acuity, color vision, and stereopsis, while higher-level vision pertains, for example, to the extraction of objects from their background, the discovery of surface geometry from a scene, and object recognition. (For simplicity, we combine the often used categories of "mid-level" and "high-level" vision into the higher-level category because, depending on one's theory of visual acquisition, the distinction is not an easy one to make when discussing visual development.)

1.2.1 Early brain plasticity in vision

For practical reasons, most studies of brain plasticity at the level of physiology have been performed in animals. This is a significant caveat: As Riesen (1947) observed in his observations of dark-reared monkeys, the duration of "apprenticeship" of visual experience required for functional usage seems to increase at higher phylogenetic levels. For example, dark-reared rats were able to make use of visual cues within 15 minutes of sight restoration in their jumping behavior and were supposedly indistinguishable from normal rats within two hours (Hebb, 1937); Riesen's monkeys required a minimum of 50 hours prior to observable "visually mediated learning"; human instances of early deprivation cataloged in von Senden (1932) "may require an even longer exposure." Although it is questionable how to compare the behavioral tasks, observational methods, and deprivation periods involved in these and subsequent studies, the general observation points out the importance of testing multiple species along the phylogenetic tree to provide for an understanding that is more likely to apply to the human visual system.

Some of the earliest and most significant studies of brain plasticity in visual cortex were those of Hubel, Wiesel and colleagues, who observed newborn kittens and the effects of experimentally-controlled visual deprivation in kittens and, later, monkeys while simultaneously measuring physiological responses from neurons in visual cortex and the lateral geniculate nucleus (LGN). In newborn kittens with almost no patterned-light experience, it was observed that the physiology of primary visual cortex (V1) resembled that of mature kittens with respect to receptive-field organization, binocular interaction, and functional architecture (line orientation and motion-direction selectivity), though the response properties of the cells were somewhat muted (Hubel & Wiesel, 1963). The presence of this visual machinery was surprising given that kittens are not able to exhibit behavioral usage of vision until perhaps 3 weeks following normal eye-opening.

Given that no experience with patterned vision was necessary for the early development of these cortical response properties, it is presumed that the rudiments of such a structure are innate. Yet, the delayed behavioral usefulness of this structure leaves open the possibility that experience is necessary to bring these cells to full maturity or to integrate it into functional vision. In order to understand the dynamics of this development, the kittens were deprived of vision in one or both eyes so that the changes in cell properties could be monitored during early development.

Wiesel and Hubel (1963) report that neurons in the LGN, a "conduit" between the optic nerve conveying the output of the retina and V1, maintained normal receptive fields despite 3 months of visual deprivation in one eye, although "marked" histological changes were observed in cells driven by the deprived eye. A very different pattern emerged for neurons in V1 (Wiesel & Hubel, 1965a). Monocular deprivation caused an almost complete inability of V1 neurons to be driven by activity from the deprived eye. Since experiments with newborn kittens showed that this wiring is indeed present from birth, this is evidence of a disruption of normal development, not simply a slowing of experience-dependent tuning. Further insight can be gained from the binocularly-deprived kittens. Although one might expect that binocular deprivation would cause a wide-scale disruption in the neuron response properties (in effect, duplicating the result in the other deprived eye as found for the one deprived eye), in fact the opposite occurred: Even without any visual input to either eye, over half of cells tested seemed "perfectly normal," although many other cells did become unresponsive.

The relatively minimal disruption in V1 response properties with binocular, as compared to monocular, deprivation seems to indicate that experience per se is not necessary for this level of development of cellular responses. Rather, some elements of this development are innate, but there is sufficient plasticity to reassign many of these properties. Figure 1.1 illustrates the remarkable extent of plasticity of ocular dominance patterns following monocular deprivation in the monkey. To further illustrate the extent of plasticity, Movshon (1976) found (in the kitten) that reversing the deprivation by suturing shut the previously unaffected eye and opening the deprived eye allowed the initially deprived eye to "regain control" of cortical neurons, though this flexibility is greatly diminished with maturational age. Recent evidence has shown, however, that in the rat a period of total visual deprivation prior to reversal can re-enable juvenile-like ocular dominance plasticity (He, Hodos, & Quinlan, 2006).

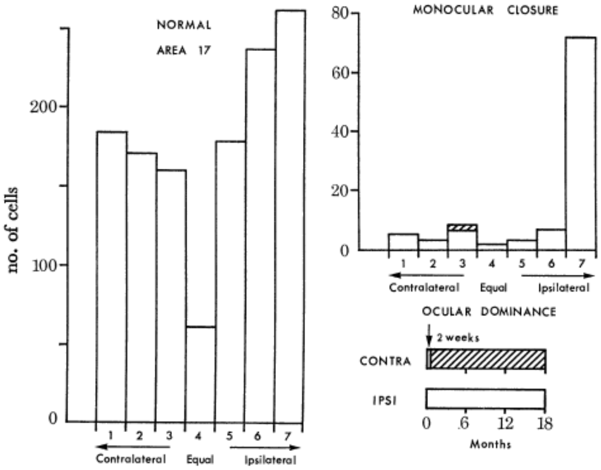


Figure 1.1 The graph on the left shows the normal histogram of ocular dominance properties of cells in V1. The graph on the right depicts the resulting pattern of ocular dominance after 18 months of monocular visual deprivation. There is a clear bias towards the non-deprived eye, illustrating the marked plasticity of the ocular dominance property of cortical cells. From (Hubel et al., 1977).

More recently, the plasticity of more advanced properties of cortical neurons has been probed. Engert and colleagues (2002) examined the response properties of neurons in the tectum (the functional analog of visual cortex in mammals) of *Xenopus* (frog) tadpoles after repetitive visual exposure to uni-directional moving bars. Tectal neurons became directionally sensitive to the direction of the moving bar as a result of the visual stimulation within minutes, and this selectivity persisted up to 50 minutes.

A more extreme form of plasticity was exhibited in the rewiring of ferret cortex to provide auditory sensory input to visual cortex and *vice versa*. Primary auditory cortex showed evidence of an orientation map responding to visual input similar to that found in V1 (though less organized) (Sharma, Angelucci, & Sur, 2000). The animals were also able to respond to light “seen” only by the auditory cortex as if it were a visual stimulus (von Melchner, Pallas, & Sur, 2000). In the human, visual deprivation in the congenitally blind can lead to activation in visual cortical areas driven by tactile or auditory stimuli (see (Burton, 2003) for a review), though it is unknown whether this activation serves a function or is merely epiphenomenal.

Taken together, these results suggest that some very elemental properties of the visual system stem from an intrinsic program, but that these properties are highly susceptible to experiential manipulation and indeed seem very flexible in their fine-grained structure and even perhaps, in the case of cortex, their locus. It would be premature, however, to generalize from these results that full, functional vision can develop in arbitrary areas of cortex with the right kind of visual stimulation, as the correspondence between these studies and the behavioral ramifications of their manipulations has not yet been established.

1.2.2 Lower-level visual development

Lower-level vision refers to those processes in the visual system that compute properties of the visual image that are not tied to the interpretation of the structure of the physical scene; rather, these properties simply reflect local image features. Such properties include acuity, color and luminance parsing, local motion extraction and stereopsis. In actuality, the distinction is not always clear, as the interpretation of a scene has been argued to affect the perception even of these seemingly clear-cut properties (Adelson, 1995; Gilchrist, 2006). In the checkershadow illusion in Figure 1.2a, for instance, it is argued that the interpretation of the dark strip as shadow enhances the illusion that patch A and patch B are of radically different luminances, whereas in fact they are the same. This explanation is dependent on a higher-level processing of the scene, and thus this effect would not be considered a pure low-level process. On the other hand, this illusion also exists without the use of shadows per se in the form of the simultaneous contrast illusion (Figure 1.2b) (Chevreul & Martel, 1860). This version gives rise to a slightly attenuated illusory effect. It is possible that the checkershadow illusion is a result of both low-level (simultaneous contrast) and higher-level (shadow-related) processes.[1] When considering the development of these abilities, it is important to differentiate between the processing of, say, luminance based on local properties versus the processing of this feature in conjunction with the overall interpretation of the scene. These processes may indeed have separate developmental trajectories. Nonetheless, most studies of low-level vision present these features to the observer in absence of higher-level confounds, allowing for an understanding of low-level visual development in isolation of these other factors.

---

[1] It may also be argued that local properties such as the shadow's penumbra may interact with the illusion, without resorting to a higher-level interpretation of the strip as a shadow, bringing us back to a low-level interpretation for the checkershadow illusion.

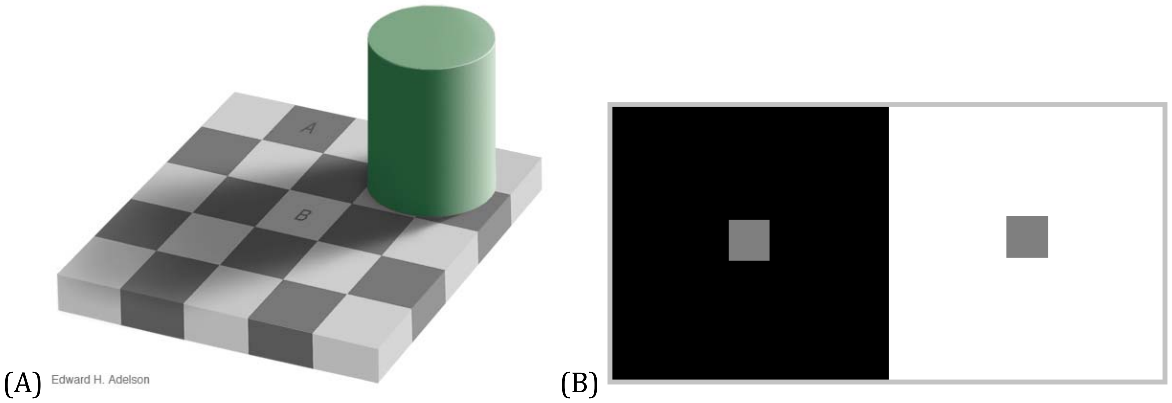


Figure 1.2 (A) The checkershadow illusion. Squares A and B have identical luminance, yet their perceived brightness is very different. The size of the effect some argue is stronger than would be expected from a difference in local contrast, thus pointing out a possible role of the interpretation of shadow (a physical interpretation) in the perception of luminance. This explanation blurs the line between low-level and higher-level vision. Figure taken from Adelson (1995). (B) The simultaneous contrast illusion illustrates a similar (though attenuated) phenomenon without the use of explicit shadows, highlighting the possibility of multiple levels of explanation.

We begin our historical perspective on low-level vision with acuity – the ability to discern change in luminance across large visual angles as well as minute distances. Acuity is of course dependent on the quality of the optical image reaching the retina due to the refractive properties and transparency of the cornea and lens, the shape of the eye, and other physical attributes of the eye itself. It is also dependent on the health of the retina and its photoreceptors to acquire the image. These considerations, however, are not our emphasis here. Instead, we consider the sensitivity of brain mechanisms to extract increasing levels of details from an image given a sufficiently clear image at the retinal level. This ability seems so fundamental to the rest of vision, that it is tempting to assume that it must be an innate property of the visual system. If this building block is not available, how could vision in any form even exist? It turns out that this view is largely correct, but with some nuances.

Normal acuity development in the human infant is relatively gradual. Figure 1.3a summarizes the progression of acuity improvement in the normal child. There is a 5-fold increase in acuity from birth to six months, and a 40-fold improvement at adult levels (Maurer & Lewis, 2001). Improvement continues until about 6 years of age. Acuity levels tend to follow gestational age (i.e., time from conception) as opposed to birth age (Weinacht, Kind, Monting, & Gottlob, 1999), indicating a prescribed timeline somewhat independent of the extra visual experience that premature infants receive compared to their normal-term counterparts.

This prolonged development is driven largely by anatomical changes in the fovea, including increases in the packing density and decreases in the diameter of the foveolar cones (Yuodelis &

Hendrickson, 1986). This evidence is consistent with a maturational account of acuity development. However, even with optimal retinal development, visual experience is necessary for the development of acuity in visual cortex: Infants with dense bilateral congenital cataracts who were tested soon after corrective surgery have much poorer functional acuity than their gestational age would predict, but improvement followed as soon as 1 week later, with further improvement still evident after 1 month (Maurer, Lewis, Brent, & Levin, 1999). Yet, improvement seems subject to a critical period. After approximately 2 months of age, a full recovery of normal acuity is highly unlikely (Figure 1.3b), though even the patient who received treatment at the age of 8 *years* attained acuity of 20/250 (Kugelberg, 1992), which qualifies as legal blindness but is still usable vision.

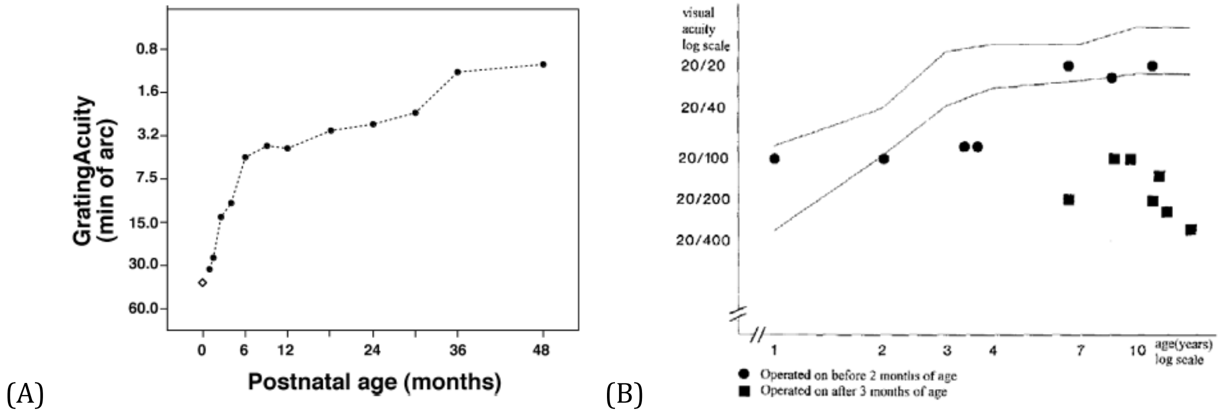


Figure 1.3 (A) Progression of grating acuity in normal children (from (Maurer & Lewis, 2001). (B) Acuity achieved in children with congenital bilateral cataracts resulting in deprivation of patterned light stimulus (Kugelberg, 1992). The horizontal scale represents the age at test. Patients wore corrective contact lenses during test. Circle points correspond to infants who were 2 months or younger at the time of corrective surgery; square points correspond to children 3 months of age to 8 years. The area between the two continuous lines represents normal age-dependent acuity.

Color discrimination in infants has also been extensively studied. Color vision, like acuity, is mostly dependent on the properties of foveolar cones. Expectedly, the development of color vision mimics the development of acuity (Kelly, Borchert, & Teller, 1997). Unlike acuity, color deprivation in the monkey during the expected critical period (3 months) has been shown to result in no diminishment in spectral sensitivity (Brenner, Schelvis, & Nuboer, 1985), indicating that color is even less influenced by cortical plasticity than acuity.

Stereopsis, another low-level visual cue, begins to develop by the age of 3 months and is relatively mature by the age of 6 months in the infant (Birch, Shimojo, & Held, 1985), although improvements can be seen until at least 5 years of age (Simons, 1981). Visual deprivation (both monocular and binocular) in congenital cataract patients generally precludes stereopsis, though some patients treated prior to 10 months of age exhibit some stereo acuity when the patients' amblyopia is compensated for (Tytla, Lewis, Maurer, & Brent, 1993). The story of "Stereo Sue"—a woman who trained herself to see in stereo after decades without the ability—is an interesting counterpoint (Sacks, 2006), but it is likely that she had the appropriate neuronal response properties already; it was the interpretation of these binocular disparity cues that may have been learned. According to Sacks, it is apparently not unheard of to lose stereopsis due to a lack of depth experience, as he observed while living in a small, windowless room in a hospital. Perhaps the perception of depth that results from the process of stereopsis, being indeed a property of the physical world rather than an image-level property, belongs in our higher-level vision category; the computation of binocular disparity would be the low-level visual process.

Finally, we come to local motion processing. The development of motion sensitivity is very rapid in the infant. Smooth-pursuit eye movements indicating the ability to track a moving target are evident in newborn infants 1-3 days old (Kremenitzer, Vaughan, Kurtzberg, & Dowling, 1979). Optokinetic nystagmus (OKN), a sign of motion coherence detection, is apparent from birth and shows little improvement over the first 6-8 months in infants, though adult thresholds are significantly better (Manny & Fern, 1990; Mason, Braddick, & Wattam-Bell, 2003). In the kitten, neurons in the primary visual cortex exhibit motion-direction selectivity at birth (Hubel & Wiesel, 1963). The immediate availability of local motion cues from birth, prior to any visual experience, points to an innate maturational provenance.

1.2.3 Higher-level visual development

Lower-level vision forms the foundation for the seemingly more difficult higher-level visual processes, yet it is unclear what aspects of lower-level vision are truly required for what we might term the "next stage" of vision. A certain level of acuity is obviously necessary, as otherwise the visual field blurs into a uniform field of luminance, but the abilities to process color and determine depth of objects through stereopsis seem useful but optional. Luminance comparison is necessary as well, while motion sensitivity is of course important for processing changes in the physical world, but it is not necessary for parsing unchanging scenes. Individuals with total color-blindness (monochromacy), as well as most other mammals, achieve otherwise normal functional vision, as do individuals with congenital stereo blindness or those with only one functioning eye; however, no known cases of individuals with a congenital case of motion blindness (Zihl, von Cramon, & Mai, 1983) are known in order to make this determination.

It is not known exactly what lower-level visual processes are necessary in order to allow higher-level processes to develop, but lower-level vision develops so rapidly that abilities, in particular acuity, should be sufficient to enable this development at one month of age or earlier. Thus, progression of higher-level visual abilities after one month should not generally be significantly deterred by the still-developing lower-level abilities.

In order to understand visual development, the most direct method is to study infants. In this vein, a highly productive line of enquiry began with Kellman and Spelke (1983) and was continued by these researchers and others. In this series of experiments, researchers studied the ability of infants to perceive an object as a whole when it is obscured by an occluder (see Chapter 4 for details). As early as 2-months of age, infants could perceive the whole object (as judged through a preferential looking paradigm) *if* the object behind the occluder underwent motion, allowing the infant to see the parts of the object segregated by the occluder moving together (Johnson & Aslin, 1995). However, the ability to perceive the object as a whole without the use of motion (using only cues based on the form of the object, such as, presumably, continuity of contour) does not develop until the age of 6.5 months (Craton, 1996), and the ability to discern the form of the hidden area did not appear until 8 months.

Although this timeline for the development of object integration processes may seem long, there is evidence that this system may not be fully developed even until adolescence. Kovács and colleagues (Kovács, Kozma, Fehér, & Benedek, 1999) found that the ability to integrate contours across large visual angles continued to improve from the ages of 5 years to 13 years. A similar pattern is observed for the development of configural face processing: improvements in configural processing are evident at 6, 8, and 10 years of age (Mondloch, Le Grand, & Maurer, 2002). The processing of global versus local patterns also changes from childhood to adolescence and even into adulthood (Scherf, Behrmann, Kimchi, & Luna, 2009), with a trend towards improvement in global integration to overcome a local processing bias.

These findings in infants and children point to a long period of visual development to attain adult-level performance, although the fundamentals of higher-level vision are available within just the first two years. Is this progression an example of learning from visual experience or a protracted maturational period? Because both processes are occurring concurrently in children, it is impossible to discern between these two possibilities.

We turn now our attention to studies of higher-level visual development involving visual deprivation. Early studies on visual deprivation from birth in animals have resulted in near-blindness immediately after sight onset, but with subsequent recovery of some visual function, for instance with rats dark-reared until maturity (Hebb, 1937) and monkeys dark-reared until 16 months (Riesen, 1947). In the case of rats, this recovery was rapid (minutes), whereas Riesen's monkeys took weeks to exhibit visual behaviors.

Historical case studies of visually deprived humans have been rare and somewhat controversial. Von Senden (1932) catalogs a number of cases of visual deprivation throughout history, including the famed and slightly better documented case of a 13 year old boy whose congenital bilateral cataracts were operated upon by William Cheselden (the first published case, from 1728). Cases of long deprivation generally result in a lack of vision, as in animals, immediately following correction, but long-term follow-up is usually given inadequate consideration. In fact, von Senden implies that the boy in the Cheselden case *did* begin to acquire functional vision: "… [H]is powers of vision only developed very gradually; it would even appear … as though he did not acquire a real understanding of the extent of space until … a year after the first [eye was operated on]" (Senden, 1960, p. 223). More recently, two important cases have attracted attention. The case of SB (Gregory & Wallace, 1963) describes a 52 year old man who obtained a corneal graft after a life of congenital blindness. SB acquired only very limited sight, but formal follow-up was limited to only six months after treatment. The case of MM (Fine et al., 2003) describes a man who lost his sight at the age of three and a half years old and received a corneal and limbal stem-cell transplant at the age of 43. MM received a more thorough examination of his visual abilities than the other patients, and the results are somewhat surprising. MM performed well on several motion tasks and some simple form tasks but failed on tasks requiring the interpretation of illusory contours or the reconstruction of depth. Even a few years of follow-up did not qualitatively alter his results. This is especially surprising given that MM had 3.5 years of normal vision prior to his loss of sight, at which point he would have had perfectly adequate vision to perform these visual tasks. This implies a significant amount of regression in his visual abilities, to the point where his recovery is poorer than that suggested by von Senden's account of Cheselden's 13 year old congenitally blind patient and certainly poorer than the cases to be discussed in our work. For now, we will leave this as a curiosity, noting only that the presence of sight followed by its disruption and subsequent recovery is a more complicated case than congenital deprivation followed by treatment; we propose a potential theory to reconcile this pattern in our final chapter.

More subtle deficiencies have been found in cases of short-term congenital deprivation, even after many years of subsequent visual experience. Children with congenital blindness due to cataracts were treated at ages between 2 and 6 months and tested, at the age of 10 years, on a task matching faces differing in subtle manipulations of facial feature configurations (Le Grand, Mondloch, Maurer, & Brent, 2001). With the relatively short period of deprivation and long period of visual experience after recovery, it is tempting to infer that the first few months of life represent a critical period for the maturation or development of full visual function, yet this conclusion seems premature given that these deficits are quite subtle. We have already discussed that configural processing develops slowly in the child, so it is less likely that this early period of deprivation would directly affect a process known to develop substantially later. Perhaps a more likely explanation is that this early period sets a bias or priority on the complement to configural information: featural processing. This bias might come from the alteration in acuity development (recall that acuity develops rapidly once vision is restored); more experience at worse acuities in the normally developing infant may favor configural processing (Pinto, Moulson, & Sinha, 2009), which may rely more heavily on lower spatial frequencies than featural processing, which requires a relatively high level of detail.

1.3 Towards a theory of visual development

From the survey above, a picture of visual development begins to emerge. Low-level visual processes can be affected by deprivation or degenerate visual input, but largely they mature on a set timeline. Where visual deprivation can significantly disrupt development (as is the case for acuity), the primary effects of the disruption occur during a critical period in the first few months of life, after which recovery may occur, but never to normal levels. Thus, we generally conclude that low-level vision is a product of maturation.

The development of higher-level vision, however, paints a more complex picture. Although deprivation in the first year of life can have ramifications for high-level processes into adolescence, these deficits are quite subtle and do not seem to qualitatively affect vision. Furthermore, these abilities emerge rather slowly over the first year of life, then continue to improve even into adulthood. Some evidence points to the existence of a critical period of several years, after which high-level vision can no longer properly emerge, but this evidence is incomplete. Studies of brain plasticity have yielded remarkable evidence of the general ability of the brain to adapt to and make use of signals intended for other cortical regions, but this evidence does not speak to the complex processes involved in high-level vision, nor does it yet address plasticity during maturity. The evidence is still insufficient to conclude whether high-level vision is acquired through visual experience or is a product of a slow maturational process.

In order to discern between these two possibilities, the timeline of maturation must somehow be separated from the process of high-level vision acquisition. That is, we require a subject population for whom visual maturation has mostly come to an end, but who have not experienced patterned visual stimulation. We have come across such a population in Project Prakash (Mandavilli, 2006), a long-term project to locate, provide treatment, and learn from the curably blind in India, which has a disproportionate number of such cases due to socioeconomic, geographical and traditional factors. Despite the relatively large number of curably blind in India, however, we seek a much smaller subset of these patients for our studies here. Our subjects must have been blind or severely compromised from birth in both eyes, and their congenital blindness must have been documented or be otherwise verifiable. Furthermore, we must have the opportunity to study each subject longitudinally, a sometimes difficult requirement in ever-changing India. Over a number of years, we have managed to find a dozen or so such cases (and a larger number of in-between cases not suitable for our discussions here), and a few of them have been chosen for publication and inclusion here.

Our first concern is whether an individual with such a history of long-term visual deprivation will ever gain normal vision. If there is indeed a hard critical period for high-level vision (after which normal vision is unattainable), then any progress following intervention may be merely a set of visual “hacks” which have little to do with normal visual development. Thus, chapter 2 presents the case of SRD, who serves to dispel the notion of a hard critical period for the acquisition of high-level vision. SRD was born with dense congenital bilateral cataracts and was by all accounts functionally blind from birth[2]. After treatment for her cataracts at the age of 12 years, SRD reportedly was unable to make significant use of her vision for about one year, after which she began to recognize objects purely by sight. We tested her extensively 20 years later and found that she had mostly normal visual skills, although with somewhat limited acuity. Hence, SRD gives us evidence that the endpoint for individuals with visual deprivation of up to 12 years is a mostly normal visual system, given adequate time for it to develop after clinical intervention[3].

---

[2] In all cases of such “blindness,” there must always be sufficient vision to enable sufficiently normal maturation of low-level visual processes. Perception of light (that is, the ability to tell light from dark but with absence of form vision) seems sufficient for this purpose.

[3] There does remain the possibility that SRD acquired a highly functional visual system through a general learning process which bears no resemblance to the normal human visual system. This, however, seems unlikely as (1) such

Knowing that even long-term visual deprivation (with the caveats of footnote 2) does not preclude the eventual acquisition of near-normal vision, our next goal was to study the stages of visual acquisition immediately following treatment. If vision is available immediately after optical correction with little exposure to visual experience, then a maturational account of visual development would almost certainly be correct. If, however, acquisition of high-level visual function requires a significant amount of time and resembles the trajectory of normal infant visual development, a theory of high-level visual development based on learning through visual experience would be a more likely alternative, as maturational processes should have already reached their endpoints prior to intervention[4]. Chapter 3 discusses the cases of three individuals who were functionally blind from birth and received treatment at the ages of 7, 13 and 29 years. They were tested immediately or soon after treatment and followed longitudinally with a large battery of tests of low- and high-level vision. To our surprise, the visual parsing performance of these subjects soon after treatment resembled the over-fragmented world observed by infants in the first few months of life. Similar to the observed performance of infants, the fragments of their visual world came together to make objects whenever these fragments shared common motion. To our greater surprise, after 8 months in the case of our youngest subject and after 18 months for our oldest subject, these individuals had begun to develop the higher-level visual abilities that allowed them to properly parse complex static scenes.

This pattern of evidence is highly consistent with a learning account of visual development, as opposed to an innate, maturational account. But how does the visual system learn such a complex process? If a visual system does not yet have a concept of “objectness,” how could this concept ever develop? Based on the pattern of performance we see in our subjects as well as the research on infant visual development, chapter 4 proposes a method of bootstrapping this learning process through the notion of common motion as a proxy for proto-object representations.

---

a visual system would likely have more significant observable differences from our expectations and (2) it would be odd and wasteful for an organism that has the capacity to create a perfectly viable visual system *de novo* to instead not use this capacity but rely on a maturational program. Thus, the most likely explanation is that SRD’s visual system is the product of delayed but mostly normal maturation, or that the visual system of typical individuals and that of SRD are acquired from visual experience through similar learning mechanisms.

[4] Although there is no physical evidence for it, there is a third alternative. Maturational processes may lie dormant, waiting for visual input in order to be re-enabled at the appropriate time. There is indeed a fourth alternative, where maturation of the higher-level visual system completes, but years of dormancy have obscured its function, relying on new stimulation to “uncover” its function, though it seems unlikely that this uncovering would result in a staged acquisition of function resembling the progression of infant development; this staged progression is indeed what we see in our subjects. Further thoughts on this complication are discussed in the final chapter of this thesis.

Finally, we conclude in the final chapter with caveats for the interpretations of our work, the implications of this research to vision science as well as the real world, and thoughts on future work to generalize our findings to crossmodal integration.

1.4 References

Adelson, E. H. (1995). Checkershadow Illusion. Retrieved July 27, 2010, from http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html

Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research*, *27*(6), 953-965. doi:10.1016/0042-6989(87)90011-3

Birch, E., Shimojo, S., & Held, R. (1985). Preferential-looking assessment of fusion and stereopsis in infants aged 1-6 months. *Invest. Ophthalmol. Vis. Sci.*, *26*(3), 366-370.

Brenner, E., Schelvis, J., & Nuboer, J. (1985). Early colour deprivation in a monkey (Macaca fascicularis). *Vision Research*, *25*(9), 1337-1339. doi:10.1016/0042-6989(85)90050-1

Burton, H. (2003). Visual cortex activity in early and late blind people. *Journal of Neuroscience*, *23*(10), 4005.

Chevreul, M. E., & Martel, C. (1860). *The principles of harmony and contrast of colours, and their applications to the arts: including painting, interior decoration, tapestries, carpets, mosaics, coloured glazing, paper-staining, calico-printing, letterpress printing, map-colouring, dress, landscape and flower gardening, etc*. H. Bohn.

Craton, L. G. (1996). The development of perceptual completion abilities: infants' perception of stationary, partially occluded objects. *Child Dev*, *67*(3), 890-904.

Engert, F., Tao, H. W., Zhang, L. I., & Poo, M. (2002). Moving visual stimuli rapidly induce direction sensitivity of developing tectal neurons. *Nature*, *419*(6906), 470-475. doi:10.1038/nature00988

Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., Wandell, B. A., et al. (2003). Long-term deprivation affects visual perception and cortex. *Nat Neurosci*, *6*(9), 915-916. doi:10.1038/nn1102

Gilchrist, A. (2006). Seeing in Black and White. *Scientific American Mind*.

Gregory, R. L. (1970). *The Intelligent Eye*. New York: McGraw-Hill.

Gregory, R. L., & Wallace, J. G. (1963). Recovery from Early Blindness: A Case Study. In *Exp. Soc. Monogr. No.2*. Cambridge: Heffers.

He, H., Hodos, W., & Quinlan, E. M. (2006). Visual Deprivation Reactivates Rapid Ocular Dominance Plasticity in Adult Visual Cortex. *J. Neurosci.*, *26*(11), 2951-2955. doi:10.1523/JNEUROSCI.5554-05.2006

Hebb, D. O. (1937). The innate organization of visual activity: I. Perception of figures by rats reared in total darkness. *Journal of Genetic Psychology*, *51*, 101–126.

Hubel, D. H., & Wiesel, T. N. (1963). Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *Journal of neurophysiology*, *26*(6), 994.

Hubel, D. H., Wiesel, T. N., & LeVay, S. (1977). Plasticity of Ocular Dominance Columns in Monkey Striate Cortex. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *278*(961), 377-409. doi:10.1098/rstb.1977.0050

Johnson, S. P., & Aslin, R. N. (1995). Perception of object unity in 2-month-old infants. *Developmental Psychology. Vol. 31(5)*, *31*(5), 739-745.

Kanizsa, G. (1979). *Organization in vision : essays on gestalt perception*. Praeger special studies. New York: Praeger.

Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognit Psychol*, *15*(4), 483-524.

Kelly, J. P., Borchert, K., & Teller, D. Y. (1997). The development of chromatic and achromatic contrast sensitivity in infancy as tested with the sweep VEP. *Vision Research*, *37*(15), 2057-2072. doi:10.1016/S0042-6989(97)00011-4

Kovács, I., Kozma, P., Fehér, Á., & Benedek, G. (1999). Late maturation of visual spatial integration in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(21), 12204 -12209.

Kremenitzer, J. P., Vaughan, H. G., Kurtzberg, D., & Dowling, K. (1979). Smooth-Pursuit Eye Movements in the Newborn Infant. *Child Development, 50*(2), 442-448.

Kugelberg, U. (1992). Visual acuity following treatment of bilateral congenital cataracts. *Documenta Ophthalmologica*, *82*(3), 211-215. doi:10.1007/BF00160767

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Neuroperception: Early visual experience and face processing. *Nature, 410*(6831), 890. doi:10.1038/35073749

Mandavilli, A. (2006). Visual neuroscience: Look and learn. *Nature, 441*(7091), 271–272.

Manny, R. E., & Fern, K. D. (1990). Motion coherence in infants. *Vision Research, 30*(9), 1319-1329. doi:10.1016/0042-6989(90)90006-7

Marr, D. (1982). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information, Henry Holt and Co. *Inc., New York, NY*.

Mason, A. J. S., Braddick, O. J., & Wattam-Bell, J. (2003). Motion coherence thresholds in infants--different tasks identify at least two distinct motion systems. *Vision Research, 43*(10), 1149-1157. doi:10.1016/S0042-6989(03)00077-4

Maurer, D., & Lewis, T. L. (2001). Visual acuity: the role of visual input in inducing postnatal change. *Clinical Neuroscience Research, 1*(4), 239-247. doi:10.1016/S1566-2772(01)00010-X

Maurer, D., Lewis, T. L., Brent, H. P., & Levin, A. V. (1999). Rapid Improvement in the Acuity of Infants After Visual Input. *Science, 286*(5437), 108-110. doi:10.1126/science.286.5437.108

von Melchner, L., Pallas, S. L., & Sur, M. (2000). Visual behaviour mediated by retinal projections directed to the auditory pathway. *Nature*, *404*(6780), 871-876. doi:10.1038/35009102

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *PERCEPTION-LONDON-*, *31*(5), 553–566.

Movshon, J. A. (1976). Reversal of the physiological effects of monocular deprivation in the kitten's visual cortex. *The Journal of Physiology*, *261*(1), 125.

Pinto, N., Moulson, M., & Sinha, P. (2009). The benefits of poor acuity for face learning. *Journal of Vision*, *9*(8), 538. doi:10.1167/9.8.538

Riesen, A. H. (1947). The development of visual perception in man and chimpanzee. *Science*, *106*(2744), 107.

Sacks, O. (2006). Stereo Sue. *The New Yorker*, *82*(8), 64–73.

Scherf, K. S., Behrmann, M., Kimchi, R., & Luna, B. (2009). Emergence of Global Shape Processing Continues Through Adolescence. *Child development*, *80*(1), 162-177. doi:10.1111/j.1467-8624.2008.01252.x

von Senden, M. (1932). *Raum-und Gestaltauffassung bei operierten Blindgeborenen vor und nach der Operation*. JA Barth.

Senden, M. V. (1960). *Space and Sight: The Perception of Space and Shape in the Congenitally Blind Before and After Operation*. London: Methuen.

Sharma, J., Angelucci, A., & Sur, M. (2000). Induction of visual orientation modules in auditory cortex. *Nature*, *404*(6780), 841-847. doi:10.1038/35009043

Simons, K. (1981). Stereoacuity Norms in Young Children. *Arch Ophthalmol*, *99*(3), 439-445. doi:10.1001/archopht.1981.03930010441010

Thelen, E., Fisher, D. M., & Ridley-Johnson, R. (1984). The relationship between physical growth and a newborn reflex*. *Infant Behavior and Development*, *7*(4), 479–493.

Tytla, M., Lewis, T. L., Maurer, D., & Brent, H. P. (1993). Stereopsis after congenital cataract. *Invest. Ophthalmol. Vis. Sci.*, *34*(5), 1767-1773.

Weinacht, S., Kind, C., Monting, J. S., & Gottlob, I. (1999). Visual development in preterm and full-term infants: a prospective masked study. *Invest. Ophthalmol. Vis. Sci.*, *40*(2), 346-353.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychological Research*, *4*(1), 301–350.

Wiesel, T. N., & Hubel, D. H. (1963). Effects of visual deprivation on morphology and physiology of cells in the cat's lateral geniculate body. *Journal of Neurophysiology*, *26*(6), 978.

Wiesel, T. N., & Hubel, D. H. (1965a). Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. *Journal of Neurophysiology*, *28*(6), 1029.

Wiesel, T. N., & Hubel, D. H. (1965b). Extent of recovery from the effects of visual deprivation in kittens. *Journal of Neurophysiology*, *28*(6), 1060-1072.

Yuodelis, C., & Hendrickson, A. (1986). A qualitative and quantitative analysis of the human fovea during development. *Vision Research*, *26*(6), 847-855. doi:10.1016/0042-6989(86)90143-4

Zelazo, P. R., Zelazo, N. A., & Kolb, S. (1972). "Walking" in the Newborn. *Science*, New Series, *176*(4032), 314-315.

Zihl, J., von Cramon, D., & Mai, N. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain: A Journal of Neurology*, *106 (Pt 2)*, 313-340.

Chapter 2. The Case of SRD: Maturation versus experience

Original title:

Vision following extended congenital blindness

Yuri Ostrovsky, Aaron Andalman, Pawan Sinha

Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, USA

This chapter is adapted from a publication which appeared in
Psychological Science 2006, vol 17(12), pp. 1009-1014.
Used by permission.

Abstract

Animal studies suggest that visual deprivation early in the lifespan can have profound consequences, causing functional blindness for the remainder of the animal's life. However, very little human data exist on this issue. Given enough time for recovery, can a person who has suffered several years of congenital blindness gain visual skills? Addressing this question is difficult since instances of sight acquisition late in life are extremely rare. We had an unusual opportunity to work with an individual in India whose case-history sheds light on this question. SRD was born blind, and remained so until the age of 12 years. She then underwent surgery for the removal of dense congenital cataracts. We evaluated her performance on an extensive battery of visual tasks 20 years after her surgery. We found that although SRD's acuity is compromised, she is highly proficient on mid/high-level visual tasks. These results suggest that the human brain retains an impressive capacity for visual learning well into late childhood. They have implications for our conceptions of cortical plasticity, and argue for treating congenital blindness even in older children.

2.1 Introduction

Through a combination of innate predispositions, cortical maturation and experience, a child comes to acquire complex visual skills. Evidence from several animal studies suggests that early visual experience is crucial for the subsequent development of visual skills (Wiesel and Hubel, 1965; Hein, Held and Gower, 1970; Bauer and Held, 1975; Hubel et al., 1977; LeVay et al., 1980). In the context of these results, an important open question is whether the unfolding of the program of visual development in humans is tied to critical periods in early childhood: Can a person acquire visual function after being deprived of sight for an extended period from birth?

To address this question, one would need to assess visual function in human subjects with various lengths of deprivation and after varying amounts of time post-deprivation. This is shown schematically in Figure 2.1. Conventional developmental studies with normal children correspond to points along the ordinate. Given that ethical considerations rule out a deliberate deprivation of sight, it has been difficult to acquire data for the rest of this space. Instances of sight onset late in the lifespan are extremely rare. Valvo estimates that fewer than 20 such cases have been reported over the past 1000 years (Valvo, 1971). Consequently, the literature on sight recovery beyond several years of age is rather sparse (Von Senden, 1932; Gregory and Wallace, 1963; Valvo, 1971; Ackroyd et al, 1974; Carlson and Hyvarinen, 1983; Carlson et al, 1986; Sacks, 1995; Le Grand et al, 2001; Fine et al, 2003; Maurer et al, 2005).

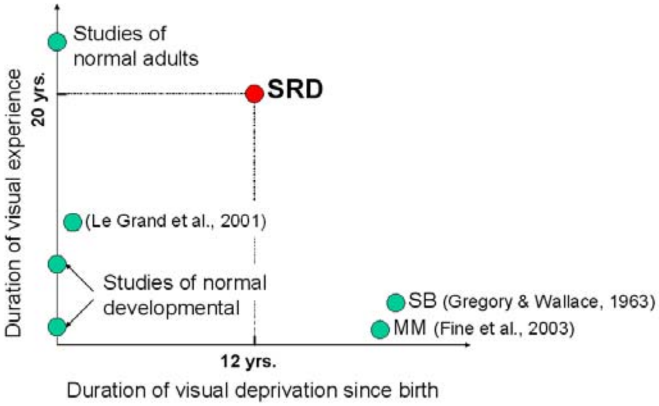


Figure 2.1 A conceptual space for studying the influence of deprivation duration and recovery duration on visual development. Previous studies have examined individuals whose histories place them close to one of the two axes. The person described in this paper, SRD, lies in a largely unexplored region of this space and allows us to assess the extent of visual recovery possible even after a very long initial period of congenital blindness, following an extended period of visual experience.

A recently launched initiative, 'Project Prakash'[1] (Sinha, 2003) has facilitated our search for individuals treated after prolonged congenital blindness. Here we report a case study from this Project: SRD, a woman who suffered an extended period of visual deprivation, was treated, and then had several years over which to acquire visual function. She corresponds to a point in the upper right sector of our two-dimensional 'Deprivation-Experience' space. SRD allows us to assess the extent of visual functionality possible with an extended recovery period, even after a very long initial period of congenital blindness.

Although Figure 2.1 places all the points corresponding to the various studies of visual skill acquisition on one plot, it is important to note that the developmental processes across the cases might be quite different from each other. Specifically, studies of late sight onset are not equivalent to studies of normal infant development. A 10 year old immediately following treatment for congenital blindness differs greatly from a newborn who has just opened her eyes. Unlike the newborn, the 10 year hold has had extensive experience of the environment through sensory

---

[1] Project Prakash is a charitable and scientific endeavor whose goal is to locate congenitally blind children in India, treat those whose blindness is correctable, and study their subsequent visual development.

modalities other than vision. This experience has likely led to the creation of internal representations that may well interact with the acquisition of visual object concepts. Gregory and Wallace (1963), for instance, found evidence that SB could rapidly transfer his haptically acquired knowledge to visual tasks. Furthermore, the deprivation may have led to structural changes in neural organization. For instance, projections from other senses may have claimed sections of the cortex that, in normal brains, would be devoted to visual processing. Thus, a priori, we cannot assume that the developmental course in the 10 year old will necessarily be similar to that in the newborn. In fact, in this paper we are particularly interested in the differences between the two cases, since they would also allow us to make inferences about neural plasticity and critical periods for specific tasks.

In this paper, we describe SRD's history and then present results from a battery of tests designed to assess her visual function. Our goal was to determine whether the prolonged deprivation that SRD had suffered had completely obliterated her ability to gain visual skills post-operatively, as a strict interpretation of the ‘critical period’ account would predict (Wiesel and Hubel, 1965; LeVay et al, 1980).

2.2 SRD: Case Description

SRD is a 34 year old female living in the Indian state of Gujarat. We met SRD serendipitously at a city eye-clinic in Ahmedabad in July, 2003. We were able to reconstruct SRD’s history from interviews with her, her parents, the physicians who attended to her at birth, and those at the clinic she visits currently.

2.2.1 Family History and Early Childhood

SRD’s father has congenital cataracts and has been blind since birth. Her mother has normal eyesight. SRD has two younger siblings: a sister, who has normal vision, and a brother with strabismus. Interviews with the midwife who delivered her indicate that SRD had profound vision problems and readily visible dense cataract formations at birth. When SRD was 1.5 months old, her parents brought her to a local physician for examination. The doctor diagnosed her as having dense bilateral congenital cataracts. Given the lack of pediatric surgical facilities, the physician recommended that surgery be postponed until SRD was older. At the age of 12 years, SRD underwent cataract removal surgery. No intra-ocular lenses were implanted. SRD is, therefore, aphakic in both eyes. Eyeglasses were provided to compensate for the missing intra-ocular lenses. The surgery succeeded in the left eye but not in the right because of complications resulting from glaucoma, providing SRD with vision only in the left eye.

Our interviews with SRD's parents were structured to help us obtain as accurate a picture about SRD's pre-operative vision as possible. The parents told us that SRD was able to tell the difference between overall levels of light, as in night versus day. However, her pattern vision was greatly compromised. She would trip over things when placed in a new environment. For this reason, SRD would never venture out of the house on her own, and was always accompanied by her father or her grandmother on her walks outside the house. Within the house (which we visited, and found to be very small – a total floor area of about 10' x 12'), SRD managed to walk around due to her familiarity with the layout of the room. The shelves and a few items of furniture were placed in fixed locations, which SRD learned through tactual experience. SRD's profound visual impairment was also evident in her interactions with people. Her parents told us that she was unable to orient towards them unless they spoke or otherwise made a sound. Thus, as best as we could assess, SRD spent the first 12 years of her childhood without visual stimuli, barring the ability to perceive overall levels of ambient illumination. However, in the absence of formal medical records, we acknowledge the necessarily approximate nature of this assessment and the possibility that her cataracts might have permitted some very limited amount of pattern vision pre-operatively. This does not significantly change the motivation for and inferences from this study, as the abilities which SRD may have had were in no way functional.

2.2.2 Post-operative Experience

Records of SRD's acuity immediately following her surgery do not exist, but she still wears the same pair of eyeglasses 20 years later (+12 dioptres, 1.5 cylindrical in the left eye). Recent tests at the Eye Clinic in Ahmedabad indicate that her prescription is appropriate, and her best corrected acuity is 20/200. After SRD's surgery, her mother explicitly taught her objects around the house. The mother claims that SRD learned to recognize her siblings and parents 6 months after surgery, and after a year, could name objects around the house purely by sight. She studied up to grade 5 in a school for blind children and then underwent two additional years of schooling in a school for normally sighted children. SRD was thus educated to the 7th grade level. She lived with her parents until she was married at the age of 27 and then moved to her husband's village.

SRD now earns money as a maid for five families in her town. She travels to the families' houses, which are 1-2 km away, on her own. She can recognize members of the families that she works for. SRD does not use navigational aids, such as a cane, but occasionally asks passersby to guide her.

We conducted a battery of experimental tests designed to probe several aspects of SRD's object perception abilities during July, 2003. The overarching goal was to determine whether these skills could be acquired even after extended visual deprivation, which in SRD's case had lasted 12 years. In order to examine several different aspects of visual skills in the limited time that we had to work with SRD, we adopted an experimental strategy that stressed breadth. The limitation of this approach is that we might have missed subtle deficits. However, it afforded us a chance to examine whether any particular skills were greatly compromised, and to assess the overall layout of her 'skill landscape'.

2.3 Experimental assessment of SRD's visual function

We decided to focus on two important domains of visual skills: basic form perception and face perception. Since these domains span a range of task complexity, it was our hope that they would be effective for assessing SRD's visual skills. The specific tests we conducted were:

Basic form perception	**Face perception**
Shape matching across space: visual search	Face/Non-face discrimination
Shape matching across time: visual memory	Face localization in natural scenes
Shape matching across transformations	Face identification

Depth perception from monocular cues

Object segmentation

Gaze estimation

Gender classification

General methods: All experiments were conducted on a laptop computer with a 14.1” screen. Screen resolution was set to 1280 x 1024 pixels. Viewing distance was approximately 30cm. In order to have a benchmark against which to compare SRD’s performance, we also recruited a normally sighted control subject, who was matched to SRD for age (32 years), gender, education (7th grade) and socio-economic status (housemaid in New Delhi). Additionally, we enlisted four subjects in the United States, ranging in age from 25 to 30 years, with normal or corrected to normal acuity. The control subject in India (the ‘status-matched control’) viewed the stimuli from a distance of 50 cm. The remaining controls (the ‘acuity-matched controls’) were seated 3 m from the screen to approximate image information loss due to SRD’s compromised acuity.

Table 1 summarizes the tests, and results from SRD and control subjects. Further details regarding experimental methods are included in the appendix to this chapter.

Table 1

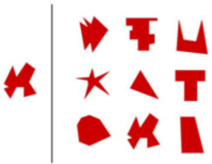
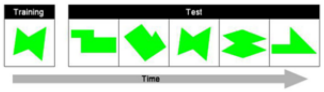
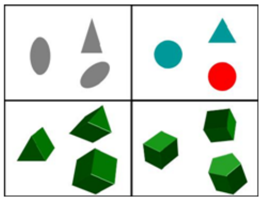
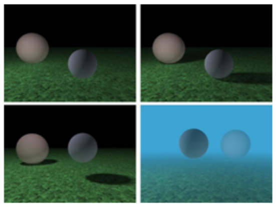
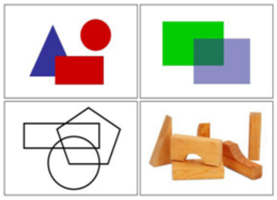
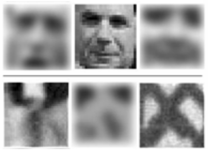
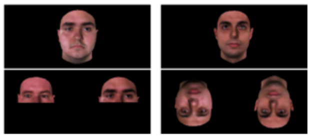
Experiment	SRD's performance	Matched control's performance	Unmatched controls' performance	Comments
Basic shape matching  Find target shape in matrix	No Errors	No Errors	No Errors	
Visual memory  Memorize target shape and find it amongst sequentially presented shapes	Hit rate: 100% (8/8) False Alarm rate: 12% (2/17)	Hit rate: 100% (8/8) False Alarm rate: 6% (1/17)	Hit rate: 92% (22/24) False Alarm rate: 2% (1/51)	SRD was well above chance ($p<0.01$; $d'>2.4$). Her performance falls in the normal range.
Matching transformed shapes  Match 2D/3D shapes while ignoring pose, scale, or photometric transformations	24/26	No Errors	No Errors	SRD was well above chance ($p<0.001$).
Depth from figural cues  Determine depth ordering using monocular figural cues	No Errors	No Errors	No Errors	
Image segmentation  Count the number of objects	No Errors	No Errors	No Errors	

Table 1 (contd.)

Experiment	SRD's performance	Matched control's performance	Unmatched controls' performance	Comments
Face/non-face discrimination  Is this a face or a non-face?	Hit rate: 15/15 False positives: 1/15	Hit rate: 15/15 False positives: 1/15	No Errors	
Face localization  Are there any faces in this scene? If so, where?	No Errors	No Errors	No Errors	
Face identification  Which of the two probe faces is the same as the target?	29/36 (See supplementary material for break-up across transformations)	No Errors	32.7/36	SRD performed well above chance ($p<0.005$), but slightly worse than controls ($p<0.05$).
Gaze estimation  In what direction is the person looking?	All responses were based on head-orientation rather than iris position	No Errors	No Errors	
Gender classification  Is this a man or a woman?	No Errors	No Errors	No Errors	

2.4 Discussion

Overall, SRD exhibited a high level of proficiency on most of our form and face perception tests. We did observe some differences in her performance relative to control subjects', such as longer reaction-times for shape matching tasks, susceptibility to overall luminance levels in face recognition and a reliance on head orientation for gaze estimation. It is possible, therefore, that SRD's strategies for image analysis might be different from those resulting from normal visual development. While this does not rule out residual impairments, this body of results does suggest that significant functional recovery is possible even after several years of congenital visual deprivation.

An important limitation we faced in studying SRD was that we were observing her visual performance well after sight onset. We had thus missed the developmental progression of her visual skills. It is possible, at least in principle, that the visual proficiencies SRD now exhibits were not learned, but rather arose through the maturational unfolding of an innately specified program. While possible, we believe that this is unlikely. The large body of research that has examined visual development in children with normal sight has underscored the importance of learning from experience in skill acquisition (Diamond and Carey, 1977; Slater, 1998; Johnson, 1998). Experiments with other individuals who have recovered sight after prolonged blindness, and who have been studied soon after sight onset, have indicated that skills are learned over time through experience with the visual environment (Gregory, 1990). Furthermore, reports from SRD's parents, though admittedly subjective, also indicate that SRD went through an extended period of learning, lasting several months, before she was able to recognize people and other objects. All of these pieces of evidence, while not definitive, make it likely that SRD's visual proficiencies were acquired via learning, instead of being innately available.

Visual learning can be thought of as comprising two components: 1. parsing the visual image, and 2. attaching labels to visual entities. Gregory and Wallace (1963) found that their subject SB showed the most improvements in his ability to attach labels to things. The early stages of SRD's post-operative experience might have been similar, but we do not have any data from that phase of her development. What our data do show, however, is that at present (18 years after her surgery) SRD's skills appear to go beyond attaching labels to familiar objects. Her ability to match novel shapes, segment images, use cues such as shading, and detect instances of objects in complex scenes, attests to her general image analysis skills.

The general picture that emerges from these experiments is an encouraging one. Our primary objective was to determine whether the long congenital deprivation would rule out even the most rudimentary of visual skills, as would be expected from a strict reading of the 'critical period' hypothesis (Wiesel and Hubel, 1965; LeVay et al, 1980). This seems not to be the case. SRD not only performed well on our tests, but can also effectively use vision for her daily activities and is now well integrated into mainstream society. These results have at least two important implications, one scientific and the other societal.

From the scientific perspective, our results suggest that the visual cortex retains its plasticity even across several years of highly compromised visual experience. This forces a rethinking of the conventional notion of developmental critical periods and also opens up some interesting questions regarding changes in cortical organization that might accompany the observed increase in visual proficiency.

From the societal health-care perspective, our results provide an argument for late stage blindness treatments. Ophthalmologists in India, as elsewhere, believe that treatment is of little use once a child is older than 7 or 8 years of age, since recovery is likely to be limited or non-existent. What SRD's results demonstrate is that even treatment at the age of 12 years can have good outcomes. Indeed, we do not know what, if any, is the upper bound on age of effective treatment. Even as Project Prakash gathers such data, we believe that it is appropriate to propose that health-care providers should not withhold treatment on the basis of age. There may well be other contra-indications to late treatments (such as disorders of the posterior eye segment), but age on its own, should not be one of them.

The one key inference we draw from these results is that the visual skill acquisition programs of the human brain remain intact to an impressive extent even when the normal timeline of their deployment has been delayed by many years.

2.5 References

Ackroyd, C., Humphrey, N. K., Warrington, E. K. (1974). Lasting effects of early blindness. A case study. *Q J Exp Psychol*, **26**: 114–24.

Bauer, J. A., and Held, R. (1975). Comparison of visually-guided reaching in normal and deprived infant monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*, **1**(4), 298-308.

Carlson S, Hyvarinen L.(1983). Visual rehabilitation after long lasting early blindness. *Acta Ophthalmol (Copenh)*, **61**(4):701-713.

Carlson, S., Hyvarinen, L. and Raninen, A. (1986). Persistent behavioural blindness after early visual deprivation and active visual rehabilitation: a case report. *British Journal of Ophthalmology*, Vol **70**, 607-611.

Diamond, R. and Carey, S. (1977). Developmental changes in the representation of faces. *Journal of Experimental Child Psychology,* **23***,* 1-22.

Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., Wandell, B. A., MacLeod, D. I. (2003). Long-term deprivation affects visual perception and cortex. *Nature Neuroscience*, **6**(9), 915-916.

Gregory, R. L. and Wallace, J. G. (1963). Recovery from early blindness: A case study. *Quarterly Journal of Psychology*. Monograph no. 2.

Gregory, R. L. (1990). Do we have to learn how to see? In *Eye and Brain*, Princeton Univ. Press.

Hein, A., Held, R., and Gower, E. C. (1970). Development and segmentation of visually-controlled movement by selective exposure during rearing. *Journal of Comparative and Physiological Psychology*, **73**, 181-187.

Hubel DH, Wiesel TN. LeVay S. (1977). Plasticity of ocular dominance columns in monkey striate cortex. *Phil Trans Soc London B.* **278**:377-409.

Johnson, S. P. (1998). Object perception and object knowledge in young infants: A view from studies of visual development. In *Percep. Development* (Ed. Slater, A.), Psychology Press: UK.

Le Grand, R. Mondloch, C. J., Maurer, D. and Brent, H. P. (2001). Early visual experience and face processing, *Nature*, **410**, 890. See also, Correction, *Nature* **412** (2001) 786.

LeVay, S., Wiesel, T.N. and Hubel, D.H. (1980). The development of ocular dominance columns in normal and visually deprived monkeys. *J. Comp. Neurol.* **191**:1-53.

Maurer, D., Lewis, T.L. and Mondloch, C.J. (2005). Missing sights: consequences for visual cognitive development, *Trends Cogn Sci,* **9,** 144-151.

Sacks, O. (1995). To see and not see. In *An Anthropologist on Mars*, Vintage Books, New York, pp 108-152.

Sinha, P. (2003) Face classification following long-term visual deprivation. *J. Vision*, **3**(9), 104a.

Slater, A. (1998). The competent infant: Innate organization and early learning in infant visual perception. In *Perceptual Development* (Ed. Slater, A.), Psychology Press: UK.

Valvo, A. (1971). *Sight restoration after long-term blindness: The problems and behavior patterns of visual rehabilitation*. New York: American Foundation for the Blind.

von Senden, M. (1932). *Space and Sight: The Perception of Space and Shape in the Congenitally Blind Before and After Operation.* Reprint, Glencoe, Illinois: Free Press, 1960.

Wiesel TN, and Hubel DH. (1965). Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. *J Neurophysiol.* **28**:1029-1040.

2.6 Appendix: Supplementary material

(Supplementary material to accompany
‘Vision Following Extended Congenital Blindness’
by Y. Ostrovsky, A. Andalman and P. Sinha)

We decided to focus on two important domains of visual skills: basic form perception and face perception. Since these domains span a range of task complexity, it was our hope that they would be effective for assessing SRD’s visual skills. The specific tests we conducted were organized as follows:

Basic form perception

- Shape matching across space: visual search
- Shape matching across time: visual memory
- Shape matching across transformations
- Depth perception from monocular cues
- Object segmentation
- Visual illusions

Face perception

- Face/Non-face discrimination
- Face localization in natural scenes
- Face identification
- Gaze estimation
- Gender classification

In what follows, we describe each of these tests and SRD's responses.

2.6.1 Basic Form Perception

Skills such as matching simple shapes, searching for them amidst distractors and compensating for appearance changes caused by image transformations, are the building blocks of complex scene-analysis and object-recognition abilities. Given their fundamental significance for spatial perception, we devoted the first half of our test battery to these basic form perception tasks.

2.6.1.1 Shape matching across space: Visual Search

We tested SRD's ability to search for designated shapes within a matrix of distractors. Three sample stimulus displays are shown in Figure 2.2. The polygonal shapes spanned, on average, 4 degrees of visual angle horizontally and vertically. The polygons were drawn in bright saturated colors, and within each trial all the polygons were the same color. The target polygons included simple regular polygons, arbitrary convex polygons, and complex polygons with numerous concavities.

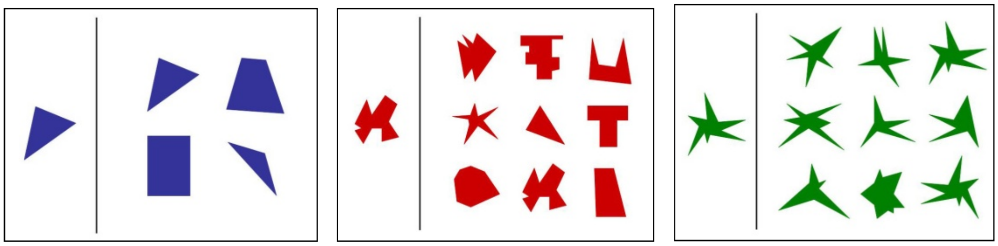


Figure 2.2 Three sample displays used to assess SRD's ability to match shapes and distinguish them from distracters. In each display, the designated target shape is shown on the left and the search matrix is one the right. SRD was asked to indicate whether the target appeared in the matrix (there were an equal number of target present/absent trials) and if so, to point to its location. SRD's performance was perfect on all trials.

Each trial proceeded as follows: The target polygon was shown alone on the left side of the screen and SRD was free to observe it for as long as she wanted. At her signal, a matrix of either 4 or 9 polygons appeared to the right of the target. SRD was asked whether the target was present within the matrix and, if so, to signal its location by touching it on the screen. SRD made no errors on all 11 trials ($p << 0.001$ compared to chance; χ^2 test), however we observed that her response times were slower than the control's. The status-matched control subject, who also made no errors, was able to correctly respond within 2 seconds on each display, while SRD's response times ranged from 5 to 20 seconds. Acuity-matched controls also performed with no errors.

2.6.1.2 Shape matching across time: Visual Memory

As the temporal analogue of spatial shape matching, we tested SRD's ability to remember polygonal shapes over short periods of time. Two sets of polygons were displayed: a training set and a test set. SRD was allowed to study the training set without time restriction (typically ranging between 5 to 10 seconds per polygon), and was then asked to identify the training set members in the test set. The test sessions followed the training after a delay of 5 seconds. The polygons were similar to those used in the visual search task. They were uniformly colored and on average spanned 4 degrees of visual angle. Within each trial, all of the polygons were the same color. We ran three trials. The first two trials each had a training set of one polygon and a test set of five polygons, while the third had three training polygons and 15 test polygons. Figure 2.3 shows the overall design of this task.

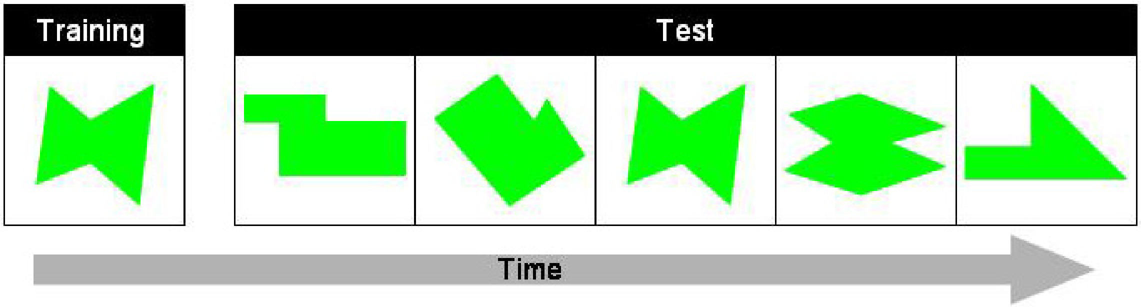


Figure 2.3 Task design for assessing SRD's ability to remember and match shapes presented across time. The training session was self-timed; SRD was allowed to look at the target pattern for as long as she wanted to. Upon presentation of the test polygons, SRD had to indicate whether they were the same as the previously seen target. SRD had a 100% hit rate and only two false positives.

SRD performed well above criterion in all three trials (hits: 8, misses: 0, correct rejections: 15, false alarms: 2; d' > 2.4[2]). The status-matched control subject performed comparably, with just one false

---

[2] Due to the absence of misses, a direct computation of the d' score is not possible. Where the number of misses is 0, we have substituted 1. Therefore, our d' measurement is an *under*-estimate of the actual d' score.

alarm (d’ > 2.8), as did the three acuity-matched controls (perfect, 2 misses, and 1 false alarm, respectively).

2.6.1.3 Shape matching across transformations

In the tests described so far, the target did not undergo any change between the training and test phases. However, in the real world, an important challenge is to recognize objects even when they may have been subjected to transformations such as rotations and scale changes. We sought to determine whether SRD could compensate for transformations during shape matching.

Each display for this task consisted of a 2D or 3D geometric figure on the left side of the screen and two similar figures on the right side. SRD’s task was to identify which figure on the right (top or bottom) matched the figure on the left, ignoring changes in size, color, orientation and (in the case of 3D shapes) lighting direction. Figure 2.4 shows a few stimulus displays.

SRD performed correctly on 24 out of 26 displays ($p << 0.001$; χ^2 test); the control subjects made no errors. (The difference in performance between SRD and the control subjects was not significant.) Notably, she answered incorrectly when discriminating between a pentagon and a hexagon whose orientations were manipulated (see upper-right panel in Figure 2.4), perhaps due to the misleading similarity in orientation of the prominent peak of the pentagon and hexagon. She was not misled by similar manipulations of simpler shapes, such as a triangle versus a square.

SRD performed well on 3D shapes, not being misled by such confounds as in the lower-right panel of Figure 2.4, where the illumination induced image-level similarities between the cube and the pentagonal prism are much greater than between the two cubes.

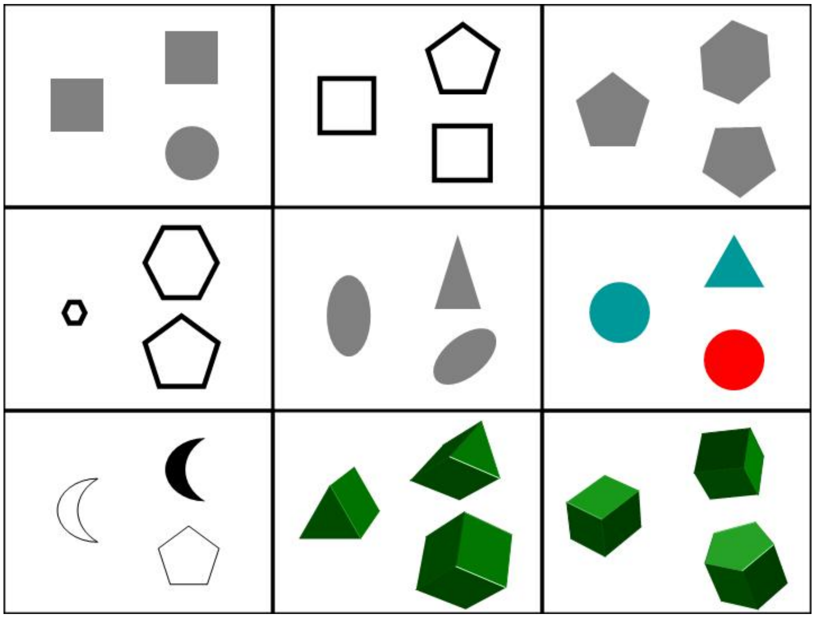


Figure 2.4 Sample displays from the shape matching across transformations experiment. For each display, SRD was asked to match the reference shape on the left with one of the two alternatives on the right, while ignoring transformations of scale, orientation, color or illumination. SRD matched 24 out of 26 trials correctly versus no errors by the control (see text for her mistakes).

2.6.1.4 Depth estimation from monocular cues

This series of 10 displays tested SRD's ability to use image-based depth cues. The basic display consisted of two spheres in a virtual world with a textured ground plane (see Figure 2.5). This was then modified through the addition of shadows, changes in image-level vertical location, changes in relative image-level size, and atmospheric haze. SRD was asked to indicate which sphere was "closer," "bigger," and "higher". In all trials, SRD's answers were consistent with a normal three-dimensional percept. That is, she appeared to be properly integrating image-based depth cues to form a 3D reconstruction of the image. She even properly used atmospheric lighting effects to determine relative nearness, despite a contradictory size-constancy cue. Responses of the control subject were identical to SRD's.

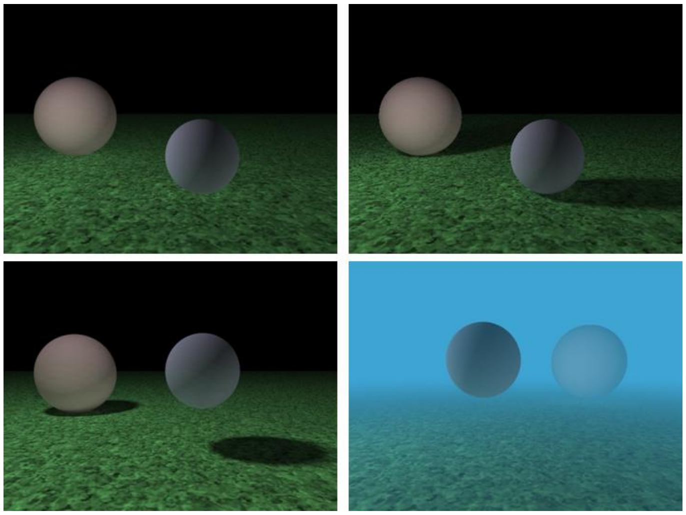


Figure 2.5 A few of the displays we used to assess SRD’s ability to estimate depth based on monocular scene cues. SRD’s assessment of all scenes was consistent with normal perception.

2.6.1.5 *Object segmentation*

So far, we have described tasks where the individual objects are entirely visible and figure-ground segregation is relatively straightforward. However, an important challenge in many settings is object segmentation – parsing a given image into coherent segments corresponding to distinct objects, even when they partially overlap and occlude each other. We tested SRD’s object segmentation ability using two tasks. The first required SRD to point to and count the number of objects she perceived within each scene. The second required her to identify the partially overlapping objects present within each scene.

For the first task, we created 20 displays each consisting of two or three polygons with or without overlap. Each polygon spanned, on average, 7 degrees of visual angle. Two special trials were included in which photographs of overlapping pool-balls and wooden building blocks were used. Figure 2.6 shows some of the experimental displays.

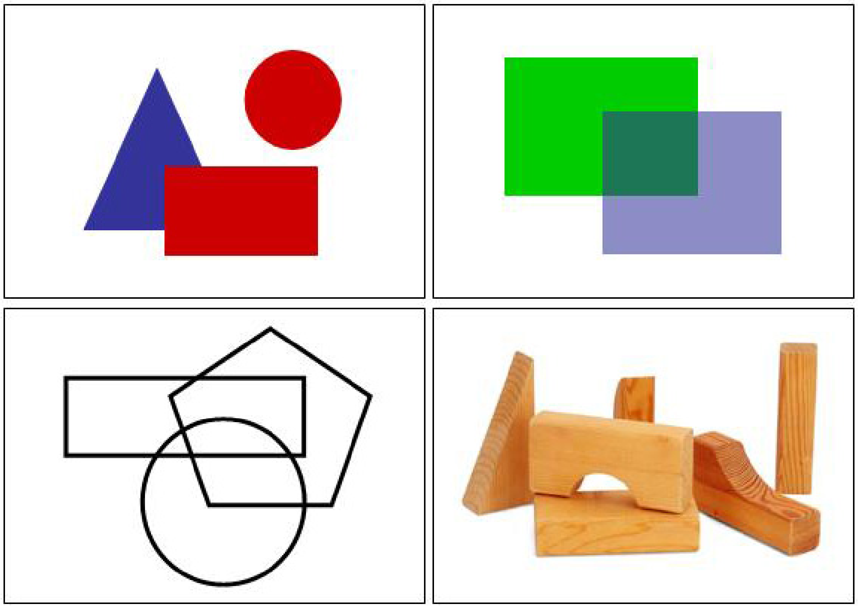


Figure 2.6 Four sample displays of the kind we used to assess SRD's image segmentation skills. For each display, SRD was asked to indicate the number of objects and point to their locations. SRD indicated the same objects as normal subjects on all trials.

SRD reported the same counts as the control in all 20 trials including those involving transparency, or unfilled line-drawings. Her success on the photographs of identically stained wooden building blocks was especially impressive given her acuity. This trial required the use of shading cues and small misalignments of bounding contours.

The second task required SRD to identify photographs of real-world objects such as trucks, cars, cups, saucers and spectacles. Each display comprised two objects with partial overlap. Figure 2.7 shows some of the experimental displays. SRD was able to correctly name all objects except a photograph of a fish, which she labeled an alligator. The control subject made no errors on either of the two tasks.

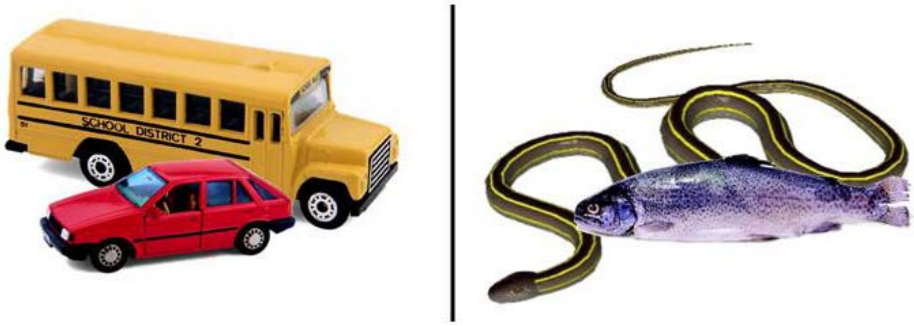


Figure 2.7 Two sample displays of the kind we used to assess SRD's object naming skills in the presence of partial overlap. For each display, SRD was asked to name and point to all of the objects in the scene. SRD identified all objects that we tested, except the fish, which was labeled as an alligator.

2.6.1.6 *Susceptibility to visual illusions*

We concluded our testing of SRD's basic form perception skills with a small battery of classic illusions. Our intent was to determine how congruent SRD's responses were to those of normal observers. Significantly divergent responses would indicate differences in the underlying visual processing mechanisms. The specific tests we conducted were the Müller-Lyer illusion, inverted T illusion, 3-dot bisection hyperacuity and the simultaneous contrast illusion. Some of our displays are shown in Figure 2.8. As described in the figure caption, SRD was susceptible to all of these illusions in a manner consistent with normal observers.

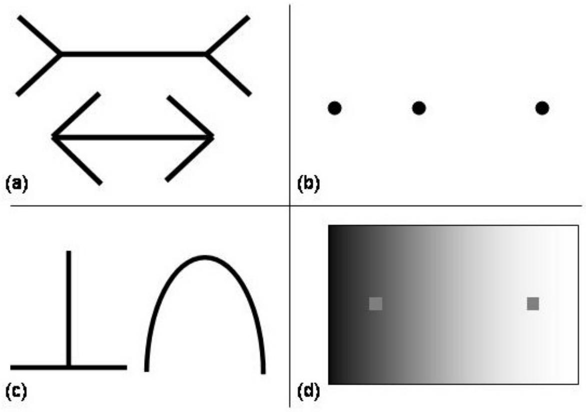


Figure 2.8 Sample displays from our tests of SRD's susceptibility to some classic visual illusions. (a) Müller-Lyer illusion; SRD was asked to indicate which line appeared longer and was instructed to pay attention to the line only, and not the overall size of the object. (b) Bisection hyperacuity display. SRD was asked to say whether the middle dot was to the left or right of center. She performed perfectly down to a 5 pixel offset (corresponding to approximately 18 minutes of visual angle). (c) Inverted T and arch illusion. SRD was asked to indicate whether the T's vertical or horizontal line segment appeared longer, and whether the arch was greater in width or height. She consistently chose the T's vertical stem and indicated that the arch

was taller than wide. (d) Simultaneous-contrast illusion. Even though the two small squares are of the same physical luminance, SRD, like normal observers, reported the square on the right as being darker.

Overall, SRD’s performance on the basic object perception tasks included in our test battery is close to that of the control subject. While this does not mean that she does not have any residual impairments, this body of results does suggest that a significant level of proficiency on form-perception skills can be acquired even after several years of congenital visual deprivation. Our next set of experiments tested SRD’s performance on more complex and ecologically relevant form-perception tasks, specifically the analysis of faces.

2.6.2 Face Processing

In the domain of face perception, we considered two broad classes of skills: (1) distinguishing between faces and non-faces (closely related to localizing faces in complex scenes), and (2) making use of intrinsic face information for assessing attributes such as identity, gender and gaze-direction. Our tests with SRD were designed to probe both of these kinds of skills.

2.6.2.1 Face/Non-face discrimination

SRD was presented with randomly interleaved grayscale face and non-face patterns and was asked to classify them as such. Images were shown one at a time and remained on the screen until SRD had responded verbally ('Yes' or 'no' to the question: Is this a face?). Our stimulus set comprised 30 patterns. Of these, 15 were faces of both genders in a frontal viewpoint, showing the face from the middle of the forehead to the chin. Each image subtended 7 degrees horizontally at a viewing distance of 30 cm. The set of non-face distracters comprised 15 patterns from natural images that had similar power-spectra as the face patterns and also false-positives of computational face-classification systems. The system we used is Rowley et al's (Rowley, H. A., Baluja, S. and Kanade, T. (1998). Neural-network based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **20**, 23-28.) scheme from CMU. Sample non-face images are shown in Figure 2.9.

Figure 2.9 A few of the face (top two rows) and non-face (bottom two rows) patterns used in our experiments. The non-face patterns comprise false alarms of computational face-detection systems and images with similar spectra as face images. SRD had a 100% hit rate, with 1 false positive. The status-matched control too had a perfect hit rate with 1 false positive.

SRD achieved a perfect hit rate, correctly classifying all of the face patterns as faces. Her only mistake was one false-positive. The status-matched control subject too had a 100% hit rate and one false-positive.

2.6.2.2 Face localization in scenes

The face/non-face classification task described above is somewhat artificial in its design. In the real-world, we are typically faced with the task of locating faces in complex scenes. Success on this task depends not only on discriminating between faces and non-faces, but also on the use of additional information arising from the rest of the body and scene structure (Cox et al., *Science*, vol 303, 115-117, 2004). We assessed SRD's face localization skills on this more naturalistic task.

SRD was shown 10 full-color natural images containing people in diverse settings. She was asked to point to faces in these scenes. The localization responses were recorded by the experimenter and their veridicality assessed relative to pre-determined face-location data for each stimulus image. Figure 2.10 shows a few sample scenes of the kind we used in the experiments. Scenes subtended 30 x 22 degrees of visual angle at a viewing distance of 30cms. The scenes included faces in different viewpoints, ranging from profile to frontal. The widths of the faces ranged from 1 degree to 10 degrees. Across the entire stimulus set, we ensured that the faces were distributed evenly over the display area so as to avoid the development of positional biases in responses.

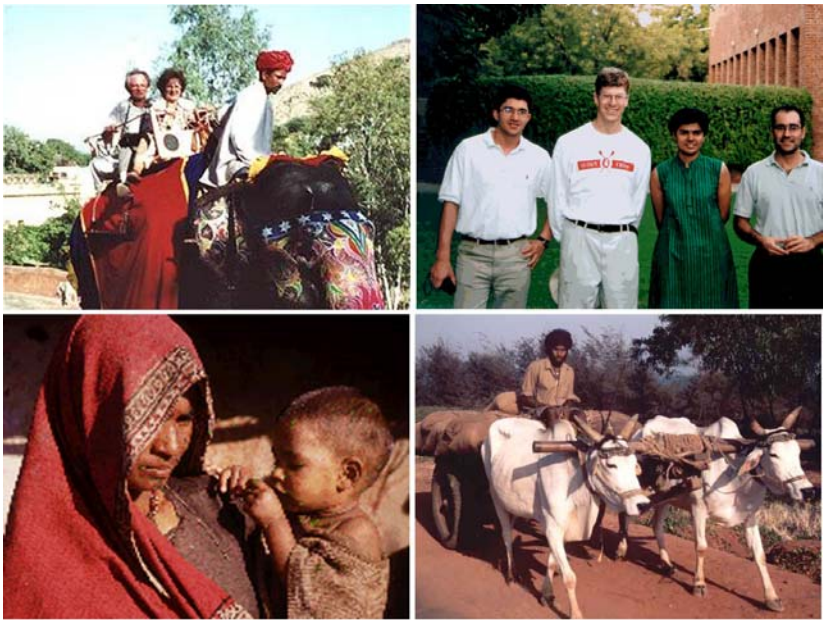


Figure 2.10 *Sample scenes of the kind that we used in our study of face localization, wherein the subject was asked to point to all faces in the scene. SRD's performance was flawless on this task.*

Both SRD and the control subject performed flawlessly on this task. SRD correctly pointed to all of the faces in the scenes and did not exhibit any false alarms.

Given SRD's impressive skills at generic face localization, we next assessed her ability to individuate people and determine their gender and gaze direction.

2.6.2.3 Face Identification

We probed SRD's face identification abilities using a set of nine face matching tasks. All nine tasks followed a delayed match to sample paradigm, and each comprised 3-5 trials. In the first seven tasks, SRD was shown a sample face to remember followed by two faces, one of the same person, the other of a different person. Both of the faces were transformed according to the task. SRD was then asked to select the face that was of the same person. The seven task transformations were: 1. vertical inversion, 2. occlusion of the upper-half, 3. occlusion of the lower-half, 4. 45 degree depth rotation, 5. Gaussian blur, 6. edge-detection (which transforms the images into black and white line drawings), 7. the luminance transform (changing the overall luminance of the face), and 8. feature replacement (eyes and the mouth in one of the test faces were replaced by the eyes and the mouth of a randomly chosen face).

The faces used in the task were chosen from the face database compiled by researchers at the Max-Planck Institute for Biological Cybernetics in Tuebingen, Germany. The faces were displayed in color and subtended 10 degrees of visual angle horizontally. They were all clean shaven and had no artifacts that could aid identification (such as eye-glasses or moles). All faces were presented on a black background and were cropped at the hairline to prevent having the hair serve as an identity cue. Figure 2.11 shows some of our stimuli.

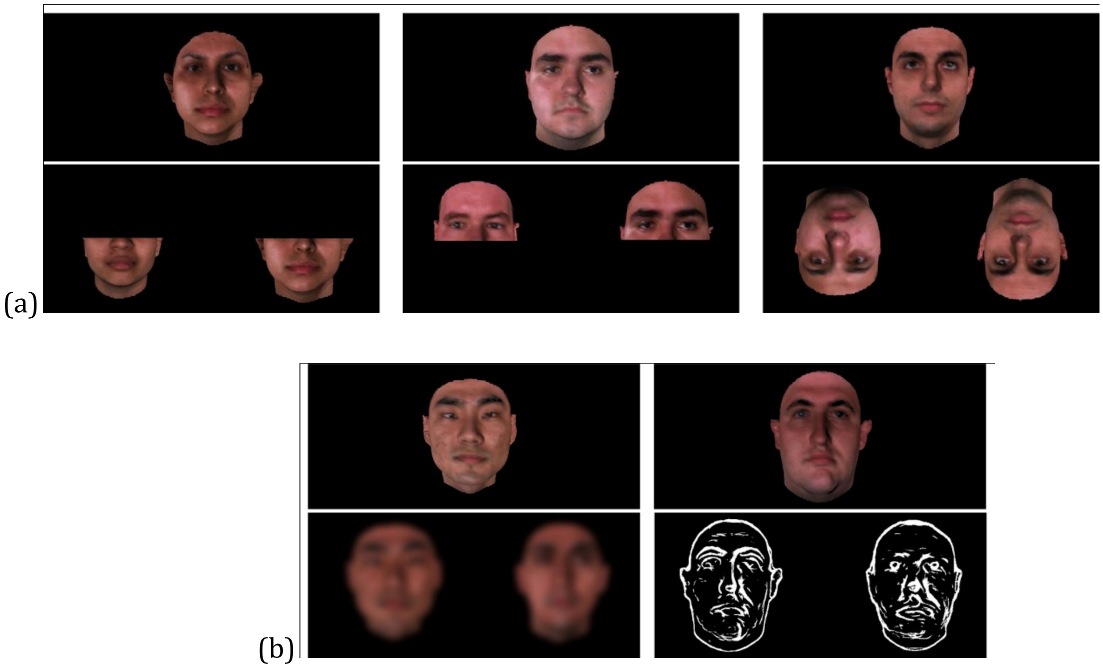


Figure 2.11 Sample stimuli used to assess SRD's face recognition performance. Shown here are five of the transformations we used – (a) occluding the upper half of the face, occluding the lower half and vertical inversion and (b) blur and edges-only. SRD was shown the target face in a self-timed fashion, and following a blank screen delay of approximately 2 seconds, the two alternatives were presented. SRD had to identify which of the two alternatives was the target face. The Gaussian blur filter we used had a radius of 14 pixels applied to face images with a pupil to pupil distance of 150 pixels. SRD's reduced acuity is akin to a Gaussian filter of approximately 4 pixel radius. The edge versions of our face stimuli involve non-linearities such as thresholding and manual noise removal, and, therefore, do not permit a simple filtering analog.

SRD's performance was as follows:

Vertical inversion: 5/5

Edges only: 3/3

Featural changes: 3/3

Upper-half occlusion: 4/5

Bottom-half occlusion: 4/5

Luminance changes: 3/5

45 degree depth rotation: 3/5

Gaussian blur: 3/5

The control subject, by comparison, performed at ceiling on all of these tests. Although we are clearly working with a limited stimulus set size, some trends are evident. SRD's performance, though not perfect, appears to be near normal for most transformations. Across the entire stimulus set, SRD's performance is significantly above chance ($p < 0.005$; χ^2 test). SRD's aggregate performance was 28/36 which was slightly worse than our acuity-matched controls at 34/36, 33/36, 31/36 and 33/36 ($p<0.05$; χ^2 test). Our status-matched control probably performed better because her displays were not adjusted for acuity, suggesting that acuity may have been responsible for at least some of SRD's remaining deficit.

2.6.2.4 Gaze estimation

In our day-to-day interpersonal exchanges, we use faces to extract more information than just identity. An important example is direction of gaze. Estimating gaze direction is a pre-requisite for tasks that require shared attention. We assessed how well SRD was able to estimate facial gaze direction. Given her impressive performance on the previous face perception tasks, we expected her to be quite proficient at determining gaze direction. We showed SRD 20 images of individuals facing the camera, but shifting their eyes to look in one of three different directions: straight ahead, left, or right (see Figure 2.12 for examples).

Contrary to our expectations, SRD performed very poorly on this task. In all cases, she reported gaze consistent with the orientation of the head and ignoring the orientation of the iris. Even slight depth rotations of the head coupled with exaggerated eye positioning in the opposite direction caused SRD to interpret gaze consistent with head pose and inconsistent with eye position. This was not due to an inability to resolve the irises in the images presented. At distances of about 1.5 m, SRD was able to correctly determine whether the pupil appeared to the left or to the right of center. When told that this positioning corresponded to gaze direction, SRD was surprised and said that she was unaware of such a correspondence. Eye-gaze estimation, thus, is a skill that is highly compromised in SRD. For the control subject, this task was trivially easy and she performed without errors on all of the stimuli. SRD's reliance on head-orientation rather than the intra-ocular iris position as an indicator of gaze direction is perhaps a consequence of her compromised acuity. Poor acuity might render iris position irresolvable at distances typical of social interactions, while still permitting head orientation to serve as a useful cue.

Figure 2.12 Sample images of the kind we used to assess SRD's eye-gaze estimation skills. It was discovered that SRD used the orientation of the head, not of the eyes, as the cue to gaze orientation. She was neither perceptually nor cognitively aware of the relationship between eye orientation and gaze direction.

2.6.2.5 Gender classification

The final test of face perception we conducted asked SRD to specify the gender of 10 faces. The 10 images were evenly divided across the genders and showed faces cropped at the forehead and the chin. Thus, hair and clothing were not available as cues. None of the women in the set were wearing the traditional (for India) red dot on their forehead. SRD correctly classified all of the images. Her response times were comparable to those of the status-matched control subject.

Chapter 3. Visual Parsing After Recovery from Blindness

Visual Parsing After Recovery from Blindness

Yuri Ostrovsky[1], Ethan Meyers[1], Suma Ganesh[2], Umang Mathur[2], Pawan Sinha[1]

[1]Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139
USA

[2]Dr. Shroff's Charity Eye Hospital
Daryaganj, New Delhi
India

This chapter is adapted from a publication which appeared in
Psychological Science (2009), vol. 20(12), pp. 1484-1491.
Used by permission.

Abstract

How the visual system comes to bind diverse image regions into whole objects is not well understood. We recently had a unique opportunity to investigate this question when we met three congenitally blind individuals in India. After providing them treatment, we studied the early stages of their visual skills. We found that prominent figural cues of grouping, such as good-continuation and junction structure, were largely ineffective for image parsing. By contrast, motion cues were of profound significance in that they enabled intra-object integration and facilitated the development of object representations which permitted recognition in static images. Following several months of visual experience, the individuals' performance improved, and they were able to use the previously ineffective static figural cues to correctly parse many static scenes. These results suggest that motion information plays a fundamental role in organizing early visual experience and that parsing skills can be acquired even late in life.

3.1 Introduction

Individuals who acquire sight late in life provide a unique window into several aspects of visual development. Such cases, however, are extremely rare; fewer than 30 have been studied in any detail over the course of the past 1000 years (Valvo, 1971). Through a concerted effort to locate such individuals in underprivileged enclaves in India, a country with an estimated 25% of the world's blind, we have been able to conduct longitudinal studies from sight onset up to 18 months later with three such patients, SK, JA and PB, providing an opportunity to add to this important but sparse body of work (Von Senden, 1932; Gregory and Wallace, 1963; Fine et al., 2003; Maurer et al., 2005).

Real-world images typically comprise many regions of different colors and luminances (Figure 3.1). Our visual systems are adept at integrating subsets of these regions into meaningful entities. How this is achieved is a fundamental question, and has been researched extensively in the domains of experimental and computational neuroscience (Wertheimer, 1938; Marr, 1982; Hummell and Biederman, 1992; Ullman, 1996; Hupe et al., 1998; Needham, 2001; Tu et al., 2003; Brady and Kersten, 2003). Much of the work has focused on the use of heuristics such as alignment of contours and similarity of texture statistics (Grossberg and Mingolla, 1985; Mumford and Shah, 1985; Field et al., 1993; Kovacs and Julesz, 1993; Leung and Malik, 1998; August et al., 1999). In circumscribed domains, these heuristics can account rather well for human performance (Koffka, 1935; Kanizsa, 1979; Elder and Zucker, 1998), but using them for analyzing real-world imagery remains an open challenge (Shi and Malik, 1997; Borenstein and Ullman, 2002). Furthermore, while it is evident that a mature visual system makes use of these cues, it is unclear whether these heuristics serve to organize information during the early stages of visual experience. Determining the nature of cues active at this time is important for elucidating the principles of visual learning and bootstrapping.

Our studies with three individuals immediately after the onset of patterned vision provided a rare opportunity to examine the bootstrapping mechanisms for visual parsing and the progression of visual abilities due to visual experience. These studies were undertaken as part of Project Prakash, our initiative in India to identify, and provide medical care to, individuals with treatable congenital blindness (Mandavilli, 2006). In working with these individuals, we also have an opportunity to examine how time-bound is the development of the parsing skills, and whether it is subject to a 'critical period'. Earlier case-reports have demonstrated that individuals who acquire sight late in life show a profound deficit in interpreting the visual confusion which they suddenly encounter (Von Senden, 1932; Gregory and Wallace, 1963; Valvo, 1971; Fine et al., 2003). These results appear to suggest that visual parsing might be subject to a critical period in the first few years of life. Despite the lack of conclusive evidence for the permanence of deprivation-induced deficits, individuals who have been blind past the age of 5 or 6 years but have treatable conditions (a situation virtually non-existent in developed nations, but which unfortunately is not as rare in the developing world) are often passed over for treatment owing to the assumed poor prognosis for recovery.

It is worth noting that working with a non-infant population provides us with both advantages and disadvantages. On the one hand, since the brain is otherwise almost fully mature, visual learning can be segregated from development of the other senses and from real-world knowledge. On the other hand, a mature brain may not undergo the same progression as an infant brain. Thus, it is more appropriate to consider this work as complementary to, rather than a replacement for, traditional studies with infants.

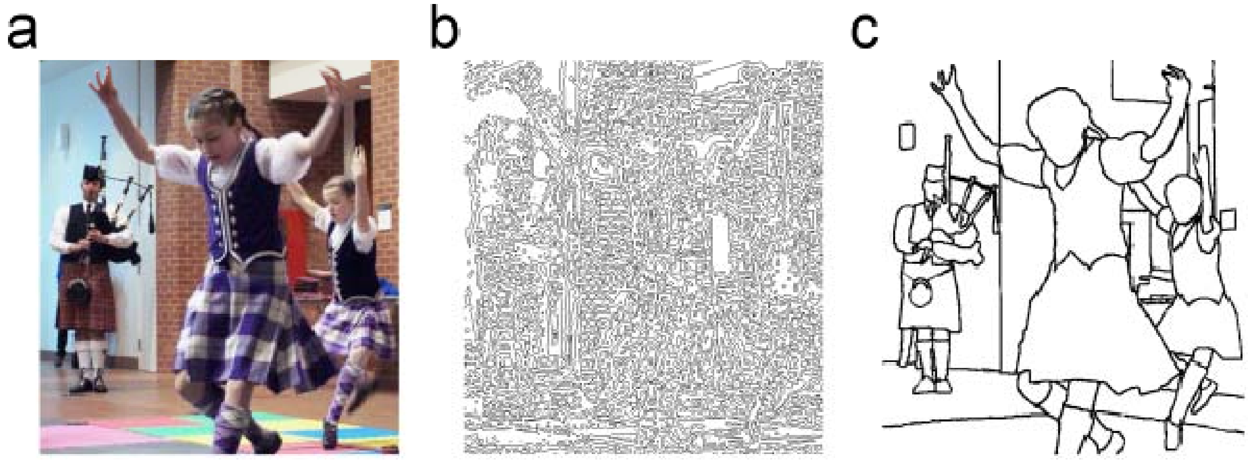


Figure 3.1 *Natural images, such as the one shown in (a) are collections of many regions of different hues and luminances, as indicated in (b). The human visual system has to accomplish the task of integrating subsets of these regions into coherent objects, as in (c).*

3.2 Methods

Participants:

SK is a 29 year old male, born in Bihar, India. By the time SK was 4 months old, members of his family noticed his inability to fixate and a lack of visually guided behaviors. Due to financial and logistical constraints, medical intervention was not sought until SK was an adolescent. At the age of 12 years, SK was examined by an ophthalmologist, who recommended surgery to correct his sight. However, the operation was cancelled due to SK's father's illness, which completely depleted his family's finances. SK was admitted to the State School for the Blind in Darbangha, Bihar, where he studied for 12 years and learned Braille. In 2000, he moved to a hostel for the blind in New Delhi and enrolled in a correspondence course, which earned him a master's degree in political science in April, 2006. It was during a visit to this hostel that the authors met SK in January, 2004.

Examinations by three independent ophthalmologists in New Delhi yielded identical assessments – SK has secondary congenital bilateral aphakia (Pratt and Richards, 1968; Johnson and Cheng, 1997), with the lenses almost completely absorbed in the anterior and posterior chambers of the right and left eye respectively. The optical pathways in the eyes are clear. SK's acuity was assessed to be 20/900. He had never been able to afford a pair of eye-glasses that could compensate for his aphakia. During our next visit to India, in July, 2004, we had SK re-examined by optometrists and ophthalmologists in New Delhi and purchased a pair of eyeglasses for him. Post-correction acuity was determined to be 20/120. The residual acuity impairment is likely due to neural amblyopia (Kiorpes and McKee, 1999).

Beginning two weeks after the refractive correction, we performed a series of experiments to assess SK's visual abilities. Tests of low-level visual function revealed SK's near-normal ability to discriminate between colors, luminances and motion directions.

In addition to SK, we also had the opportunity to work with two male children, PB and JA, whom we studied 1 month and 3 months after surgery, respectively, to correct their dense bilateral congenital cataracts. PB received treatment at the age of 7 years and JA at the age of 13 years. PB was born in a village near Panipat, Haryana. His family has a long history of congenital blindness. Both PB and his elder sister TB (age 12) were congenitally blind, as were his father, paternal grandmother, great-grandmother, two aunts, and an uncle. PB has been enrolled in the Blind Relief Association's school in Delhi since the age of 4½ years. The parents did not pursue treatment for PB and TB because a doctor incorrectly pronounced TB's condition to be untreatable due to the development of nystagmus. A botched eye surgery that PB's uncle had undergone a few years earlier further dampened the parents' desire to seek treatment for PB and TB. We came across PB in an outreach eye screening session we had organized in his school. His condition (congenital bilateral cataracts) was determined to be treatable. In December 2005, PB underwent a small-incision cataract surgery with intra-ocular lens implantation in both eyes, improving his acuity from light-perception to 20/100.

JA was born in Bijnor, Uttar Pradesh. He has five siblings, three sighted (ages 21, 19, and 8) and two congenitally blind (ages 17 and 7). Both parents are illiterate, and JA has never received any education. JA received cataract surgery and an intraocular lens implant at the Shroff Charity Eye Hospital in both eyes (Right Eye: September, 2005; Left Eye: October, 2005) to treat his dense bilateral congenital cataracts, resulting in an improvement in his acuity from light-perception to 20/80.

In what follows, we describe results from all three individuals. Practical constraints allowed us to work with SK more thoroughly than with PB and JA, preventing us from replicating every experiment from the battery with the children. A few of the following results graphs, therefore, show only SK's data. For convenience, we shall refer to SK, JA and PB collectively as the 'Recently Treated' group.

SK volunteered his participation and was not paid, other than being compensated for transportation costs. The families of PB and JA were compensated for their expenses. Subjects were free to take as many rest breaks as they wished during the course of the testing. We also enlisted four normally sighted adult control subjects. These subjects came from a similar social tier as our experimental group and had received a basic education through high school.

Procedures:

Our visual parsing studies comprised seven experiments, which assessed the subjects' responses with images of simple shapes. Their task was to say how many objects there were in each of the images, point to where they were, and whenever possible, name them. (All subjects were already familiar with common shape names through touch.) Figure 3.2(a) shows the specific tasks and representative stimuli corresponding to these experiments. Each of the conditions in the experiments to probe integration skills comprised 10 distinct trials. The recently-treated subjects' viewing distance averaged 40 cm. Control subjects' viewing distance was scaled to simulate image information loss due to the reduced acuity of the recently-treated group. All stimuli were presented until a response was given.

3.3 Results

The graph in Figure 3.2(b) shows the recently-treated subjects' performance on these tasks relative to that of the control subjects. Their responses on this experimental battery exhibited a consistent pattern. They had no difficulty in enumerating individual geometric shapes when presented by themselves, or even in the presence of other shapes, so long as they were non-overlapping (experiment 'A'). However, with overlapped figures, presented either as line-drawings or as filled transparent surfaces (experiments 'B' and 'C'), the recently-treated subjects' responses were very different from controls'. They perceived all closed loops and regions of uniform luminance as distinct objects. All errors we observed were such errors of 'over-fragmentation'. Thus, for instance, when viewing two overlapping squares, the recently treated subjects would invariably parse it as three objects. Using lines of different colors or luminances as potential aids to segmenting the component objects did not change this pattern of results. To ensure that the subjects understood the task, we had told them at the start of the experiment that figures may be overlapped (a notion they were familiar with from prior haptic experience) and that they had to indicate the number of 'objects' rather than 'regions'.

With opaque overlapping shapes (experiment 'D'), SK was able to correctly indicate the number, but was at chance at determining their depth-ordering (experiment 'E'). Extended contours made up of a series of separated line segments, embedded in a field of randomly oriented ones (experiment 'F') were only infrequently detected by the recently-treated subjects. In images of three-dimensional shapes, such as cubes or pyramids (experiment 'G'), with surfaces of different luminance consistent with lighting and shadows, the recently-treated subjects reported perceiving multiple objects, one corresponding to each facet. They were unable to integrate the facets into the percept of a single 3-D object.

In summary, the recently-treated subjects' performance with this stimulus set indicated a profound inability to use cues of contour continuation, junction structure and figural symmetry to analyze the images presented. The subjects' tendency to perceive the aforementioned stimuli in a fragmented manner was also reflected in their tracings of simple figures. Figure 3.2(c) shows a tracing by SK (in red) over a drawing by one of the authors.

Next we investigated the functional significance of the recently-treated subjects' atypical image parsing skills. Given their pronounced tendency to over-fragment images, we reasoned that their ability to veridically segment and recognize real-world images would be compromised. To test this hypothesis, we assessed their naming performance on a set of 50 images of common objects. SK was able to recognize only 26% of all images shown to him, JA recognized 34% and PB only 18%. We asked subjects to point to objects in these images and also to indicate their extent, even if they could not name the objects. We found that their responses were driven by low-level image attributes; they pointed to regions of different hues and luminances as distinct objects. This approach greatly over-segmented the scenes and partitioned them into meaningless regions, which are unstable across different views and uninformative regarding object identity. A robust object representation is difficult to construct based on such fragments. Figure 3.2(d), which shows SK's responses on three sample images, illustrates this tendency towards over-fragmentation. In separate computational simulations, we have found that the subjects' parsing can be largely accounted for by a simple computational algorithm of luminance and hue-based segmentation (Figure 3.2(d), lower row).

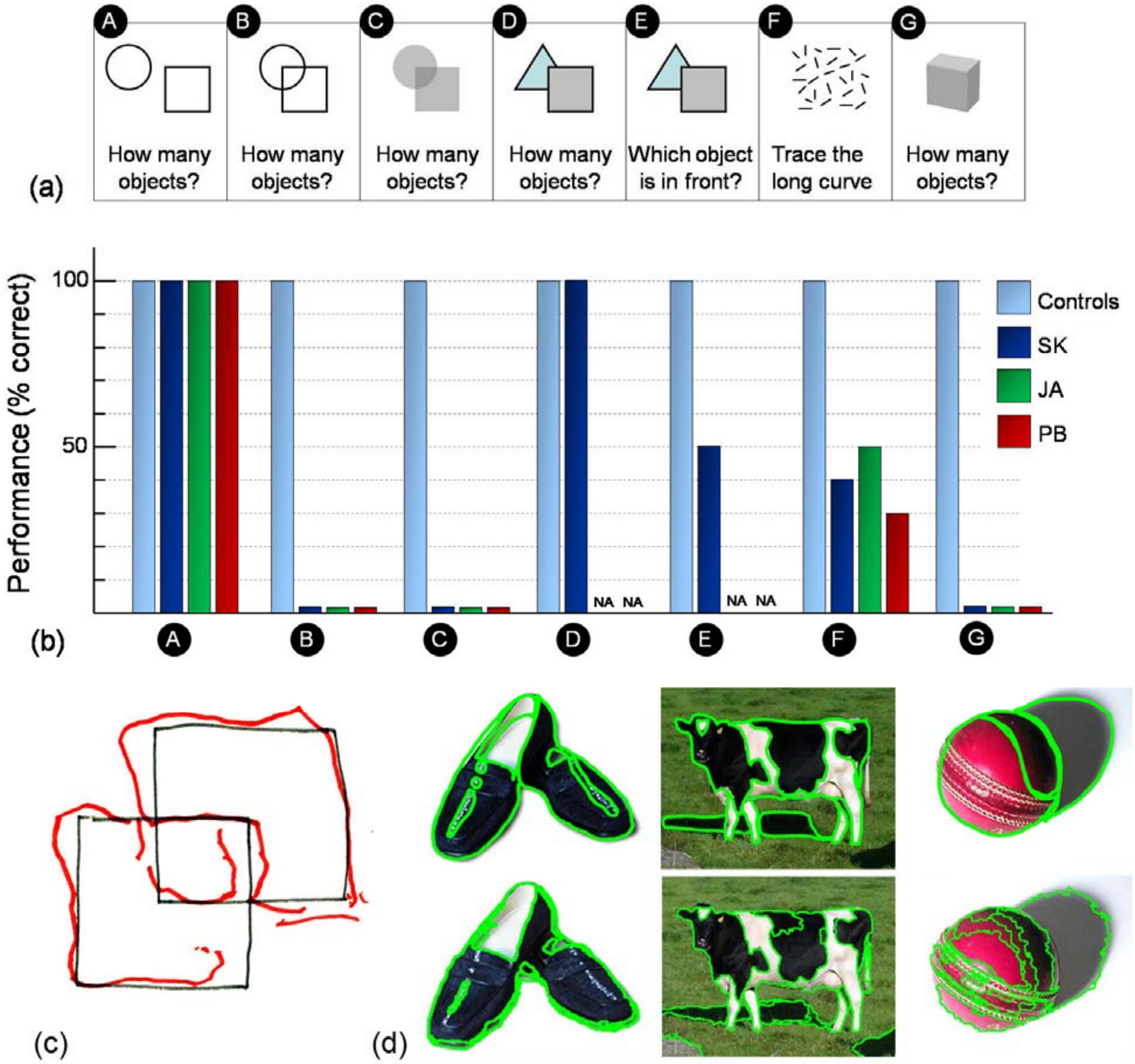


Figure 3.2 Subjects' parsing of static images. Seven tasks (a) were used to assess the recently treated subjects' ability to perform simple image segmentation and shape analysis. The graph (b) shows the performance of these subjects relative to the control subjects on these tasks. "NA" indicates that data are not available for a subject. S.K.'s tracing of a pattern drawn by one of the authors (c) illustrates the fragmented percepts of the recently treated subjects. In the upper row of (d), the outlines indicate the regions of real-world images that S.K. saw as distinct objects soon after sight recovery. He was unable to recognize any of these images. For comparison, the lower row of (d) shows the segmentation of the same images according to a simple algorithm that agglomerated spatially adjacent regions that satisfied a threshold criterion of similarity in their hue and luminance attributes.

So far, we have described the recently-treated subjects' performance exclusively with static imagery. In order to make our experiments more representative of everyday visual experience, which typically involves dynamic inputs, we created a new set of stimuli that incorporated motion cues (Figure 3.3). The recently-treated subjects' task here was the same as before – to indicate the number of objects shown. The individual shapes underwent independent smooth translational motion. For overlapping figures, the extent of movement was such as to maintain an overlap at all times.

The inclusion of motion brought about a dramatic change in the recently-treated subjects' responses. As data in Figure 3.3 indicate, they were able to produce correct responses for a majority of the dynamic stimuli. Motion also allowed SK to perceive shapes embedded in noise, a task that he was entirely unable to do in the static condition. Motion thus appears to be instrumental for enabling the subjects to link together parts of an object, and segregate them from the background.

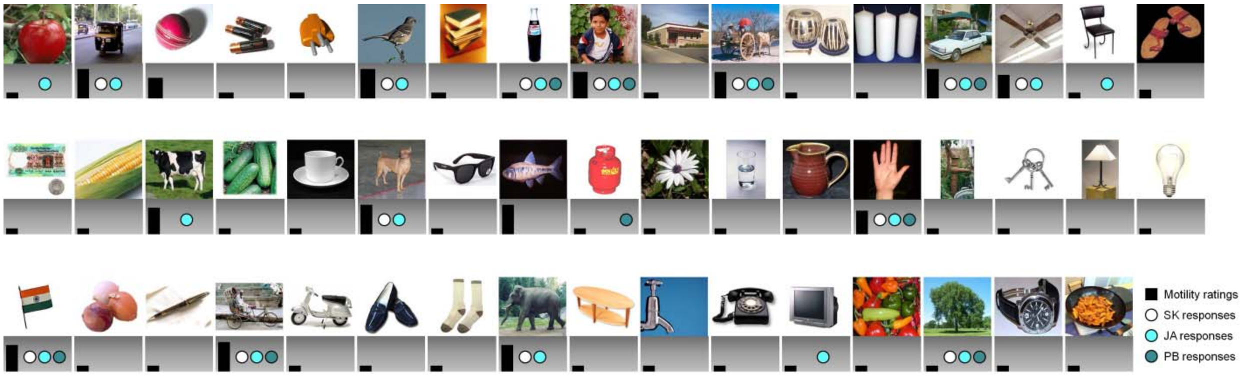
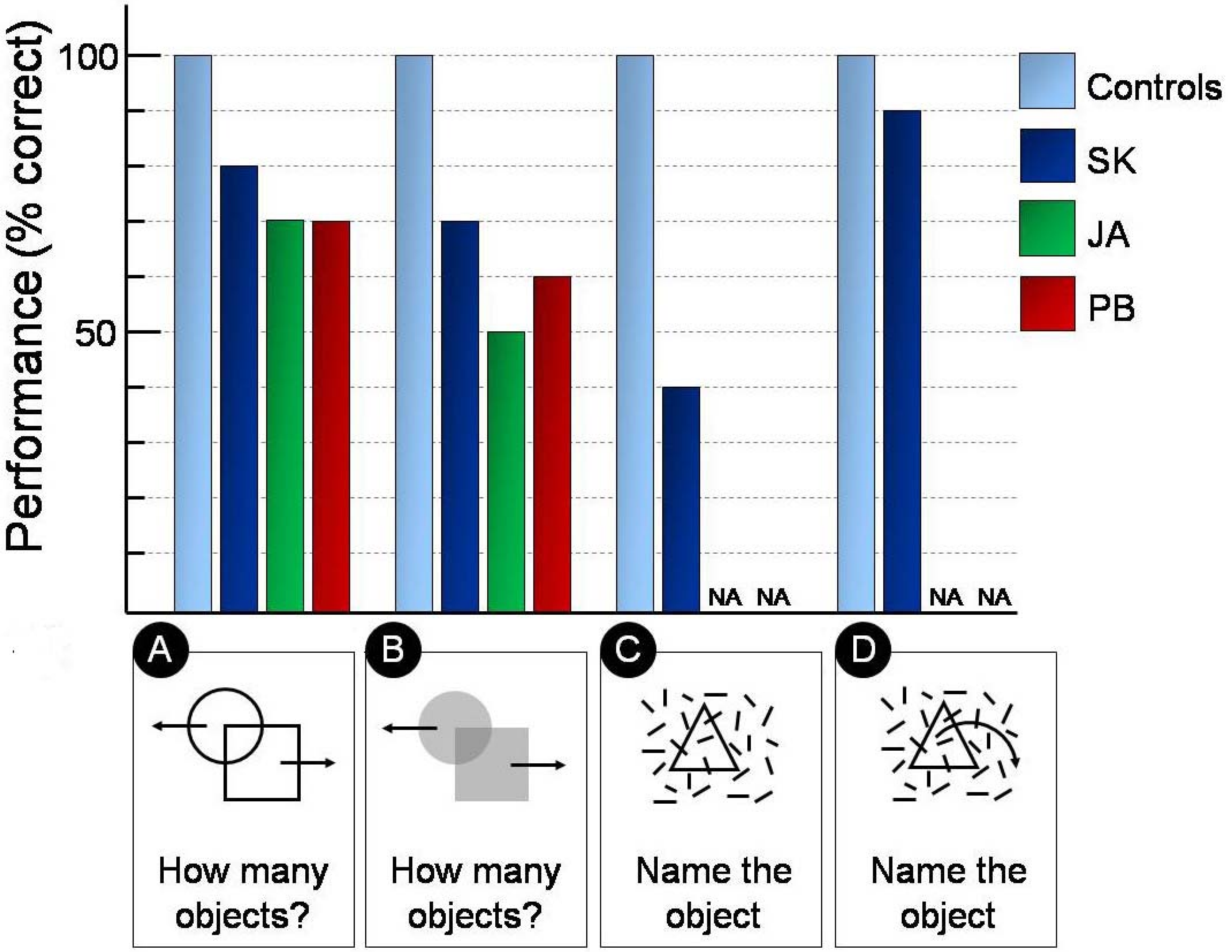


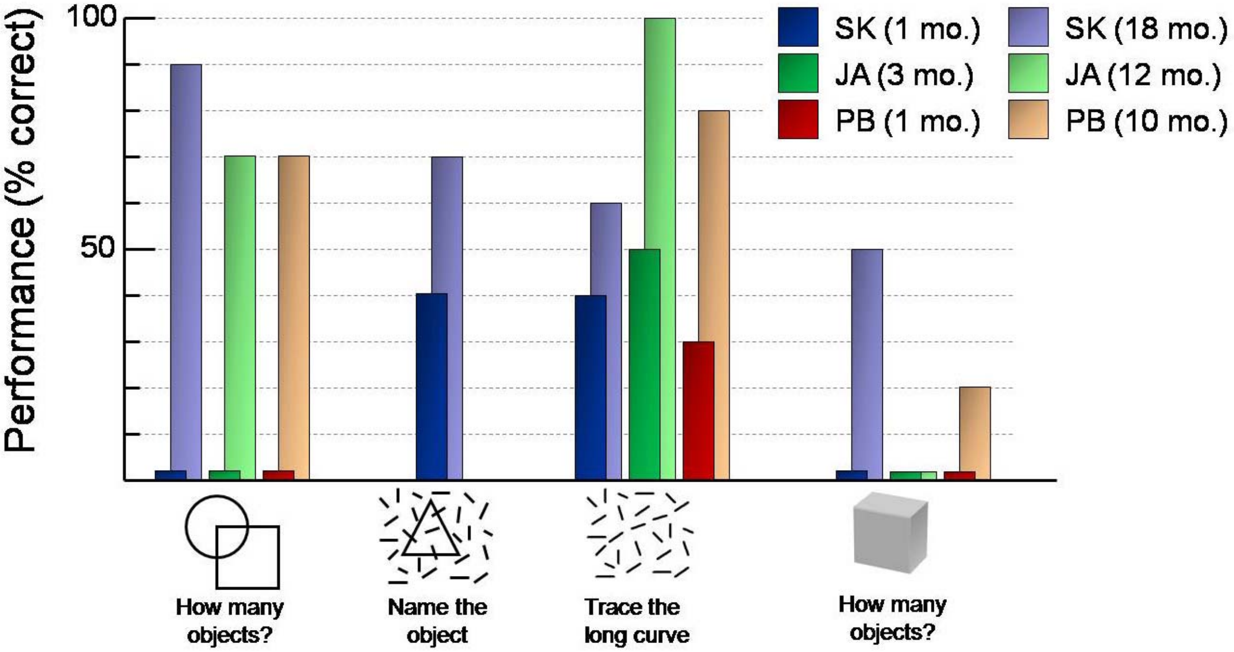
Figure 3.3 Motility ratings of the 50 images used to test object recognition and the recently treated subjects' ability to recognize these images. Motility ratings were obtained from 5 normally sighted respondents who were naive as to the purpose of the experiment; the height of the black bar below each object indicates that object's average rating on a scale from 1 (very unlikely to be seen in motion) to 5 (very likely to be seen in motion). The circles indicate correct recognition responses.

The recently-treated subjects' recognition results with real-world images, alluded to above, provide evidence of another role that motion might play in their object perception skills. In examining which images the recently-treated subjects were able to recognize, an interesting pattern became evident. As shown in Figure 3.4, the recently-treated subjects' recognition responses show a significant congruence with independently derived motility ratings of objects shown in the images ($p < 0.01$; χ^2 test for each of the three subjects). A plausible, though not definitive, explanation of the congruence between the recently-treated subjects' responses and ratings of motility, is that motion of objects helps bind their constituent regions into cohesive representations, which can then be used to recognize instances in new inputs that may well be static. It appears, however, that motion information is not used to the exclusion of figural cues, since preliminary tests with point-light walkers of the kind devised by Johansson (1973) proved ineffective for conveying the impression of a person. None of the three subjects in the recently treated group was able to perceive such displays as anything other than a collection of moving dots.



3.4 Recently treated and control subjects' performance on tasks designed to assess the role of dynamic information in object segregation. "NA" indicates that data are not available for a subject. In the illustrations of the four tasks, the arrows indicate the path of movement.

We examined changes in the recently-treated subjects' performance as a function of time after treatment. SK's performance pattern was unaltered when we tested him 6 and 12 months after treatment. Given the relatively mature age at which he had received treatment, we were not hopeful of observing much visual recovery. However, the follow-up tests we conducted at 18 months post-treatment demonstrated that SK's visual skills, while still not normal, had registered a significant improvement. The results are summarized in Figure 3.5. Essentially, where SK previously required motion cues to integrate regions into objects, he could now perform the same tasks in static images. As Figure 3.4 shows, PB and JA too exhibited a similar improvement in their ability to parse static images as SK, when tested several months after initial treatment (10 months for PB and 12 months for JA).



3.5 The recently treated subjects' performance on four tasks with static displays soon after treatment and at follow-up testing after the passage of several months (indicated in the key).

3.4 Discussion

Taken together, these results provide us a longitudinal glimpse into the development of visual parsing skills later in life. They suggest that the early stages of this process are characterized by integrative impairments. These impairments lead to perceptual over-fragmentation of images, and thus compromise recognition performance. However, the use of motion information effectively mitigates the integrative difficulties. During the early stages of visual learning, motion appears to be instrumental both in segregating objects and also in binding their constituents into representations for recognition.

We derive confidence in the generality of the results described here based on the consistency between the three subjects, and also their congruence with findings from previously reported case-studies of sight-recovery. While the earlier papers on sight recovery in adulthood (Gregory and Wallace, 1963; Fine et al., 2003) do not specifically focus on the individuals' region integration skills, they report difficulties consistent with impairments in these skills during natural image recognition (for instance, Gregory and Wallace (1963, page 24), in describing their patient SB, wrote: "We formed the impression that he [SB] saw [the natural scenes as] little more than patches of colour.") and in simple image parsing (Fine et al., in their report on MM (2003, page 915) wrote: "[MM] described two overlapping transparent squares as three surfaces with the central square in front."). Furthermore, in these past cases, as in the present one, motion sensitivity was evident soon after treatment.

The privileged status of motion observed with our recently-treated individuals is reminiscent of results reported in the infant literature. Although infants eventually come to be able to use static figural cues for object segregation (Needham, 1998, 1999), segmentation from motion arises at least two months prior to the ability to segment from static cues (Arterberry and Yonas, 2000; Johnson, 2003), and their ability to link spatially separated parts of a partially occluded object is initially driven strongly by common motion (Kellman and Spelke, 1983; Johnson et al., 2002). It is interesting to find this point of overlap between two developmental progressions that are separated by several years, given that maturational processes of the older group would presumably have already completed their time course. The neural underpinnings of this similarity are unclear. Its perceptual utility, however, admits a conjecture. It is possible that the early availability of motion sensitivity in the primate brain (Kiorpes and Movshon, 2003, 2004) serves an adaptive purpose by providing a scaffolding for acquiring static figural analysis skills. By observing the correlations between motion-based groupings and static cues, such as aligned contours, the visual system might learn to use the latter by themselves as proxies for grouping (Cavanagh, 1993). This conjecture regarding potential dependencies between early and later developing visual skills has significant implications for theoretical models of visual learning (Sinha et al., in press).

Our experimental results complement past studies of visual development in infancy and after sight restoration. First, they provide evidence that region integration via figural cues is unlikely to be merely a maturational process, unfolding with age, but rather a visually driven developmental one. Second, they highlight the limited grouping efficacy of static figural cues, such as spatially contiguous collinear contours, early in the visual-learning timeline. These cues have conventionally been assumed to be of fundamental significance for spatial integration (Wertheimer, 1938; Ullman, 1996; Sigman et al., 2001). Third, the results provide evidence that motion cues might facilitate the building of linked assemblies of regions that can serve as representations for recognition in new inputs. In this way, they serve to connect basic grouping phenomena to real-world object recognition. Overall, these results suggest that dynamic information provides a key organizing influence for early visual processing.

The evidence of marked improvement in our subjects' performance over the course of several months suggests that visual skills related to the complex task of image parsing can be acquired even after a prolonged delay, although the rate of acquisition slows down with age, possibly due to decreases in plasticity. Furthermore, the subjects' visual experience during this period derived from their normal daily activities; no special training was provided. Indeed SK resided at a hostel for the blind with no sighted residents to provide instruction. These results, along with a case we have reported previously (Ostrovsky et al., 2006), argue for a tempering of the strict critical-period idea, and provide cause for optimism for the many blind individuals who are candidates for treatment. The human brain, it appears, retains at least some measure of its ability to launch programs of visual learning even after extended periods of visual deprivation. Furthermore, the insights into the progression of visual skill acquisition gathered from these studies point to possible rehabilitative programs for the often overlooked patients with congenital sight deprivation.

Acknowledgments

The authors wish to thank SK, PB and JA and their families, the staff at New Delhi's Shroff Charity Eye Hospital and Profs. Richard Held, Beatrice de Gelder and Scott Johnson. Funding for this work was provided by NIH grant R21 EY015521, the Alfred P. Sloan Foundation, and the John Merck Scholars Fund.

3.5 References

Arterberry, M. E., and Yonas A. (2000). Perception of three-dimensional shape specified by optic flow by 8-week-old infants. *Perception & Psychophysics*, **62**(3), 550-556.

August, J., Siddiqi, K., and Zucker, S.W., (1999). Contour Fragment Grouping and Shared, Simple Occluders, *Computer Vision and Image Understanding*, **76**(2): 146-162.

Borenstein, E. and Ullman, S. (2002). Class-specific, top-down segmentation. A. Heyden et al. (Eds.): Springer Verlag, *LNCS 2351*, 109–122.

Brady, M. J., & Kersten, D. (2003). Bootstrapped learning of novel objects. *Journal of Vision*, **3**(6), 413-422, http://journalofvision.org/3/6/2/, doi:10.1167/3.6.2.

Cavanagh, P. (1993). The perception of form and motion. *Current Opinion in Neurobiology*, **3**, 177-182.

Elder, J., and Zucker, S.W. (1998). Evidence for boundary-specific grouping in human vision, *Vision Research*, **38**(1): 143-152, 1998.

Field, D., Hayes, A., Hess, R. (1993). Contour integration by the human visual system: evidence for a local "Association Field. *Vision Res.* Vol. **33**, No 2 pp: 173-193.

Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., Wandell, B. A., MacLeod, D. I. (2003). Long-term deprivation affects visual perception and cortex. *Nature Neuroscience*, **6**(9), 915-916.

Gregory, R. L. and Wallace, J. G. (1963). Recovery from early blindness: A case study. *Quarterly Journal of Psychology*. Monograph no. 2.

Grossberg, S. and Mingolla, E. (1985). Neural Dynamics of Perceptual Grouping: Textures, boundaries and emergent segmentations, *Perception and Psychophysics*, **38**(2):141--171.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, **99**, 480-517.

Hupe, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P. and Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature*, **394**, 784-787.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, **14**, 201-211.

Johnson, B. L., and Cheng, K. P. (1997). Congenital aphakia: a clinicopathologic report of three cases. *J. Pediatr. Ophthalmol. Strabismus.* **34**(1):35-39.

Johnson, S. P., Bremner, J. G., Slater, A., Mason, U., & Foster, K. (2002). Young infants' perception of unity and form in occlusion displays. *Journal of Experimental Child Psychology*, ***81***, 358-374.

Johnson, S. P. (2003). Development of fragmented vs. holistic object perception. In G. Schwarzer & H. Leder (Eds.), *The development of face processing,* 3-17, Cambridge, MA: Hogrefe & Huber.

Kanizsa, G. (1979). *Organization in Vision: Essays on Gestalt Perception*, New York: Praeger.

Kellman, P. J. and Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology*, **15**, 483-524.

Kiorpes, L. and McKee, S.P. (1999). Neural mechanisms underlying amblyopia. *Current Opinion in Neurobiology,* **9**: 480-486.

Kiorpes, L. and Movshon, J.A. (2003). Neural limitations on visual development in primates. In *The Visual Neurosciences*, ed. Chalupa, L.M. & Werner, J.S., pp. 159–173. Cambridge, Massachusetts: MIT Press.

Kiorpes, L. and Movshon, J.A. (2004). Development of sensitivity to visual motion in macaque monkeys. *Visual Neuroscience* **21**(6): 851-859.

Koffka, K. (1935). *Principles of Gestalt Psychology*, New York: Harcourt, Brace and World.

Kovacs, I. and Julesz, B. (1993). A Closed Curve is Much More than an Incomplete One: Effect of Closure in Figure-Ground Segmentation. *Proc. Nat. Acad. Sci., USA*, **90**:7495-7497.

Leung, T. and Malik, J. (1998). Contour continuity in region based image segmentation. In *Fifth Euro. Conf. Computer Vision*, Freiburg, Germany.

Mandavilli, A. (2006). Look and Learn. *Nature,* 441, 271-272.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* W. H. Freeman and Company: New York.

Maurer, D., Lewis, T. L., and Mondloch, C. J. (2005). Missing sights: consequences for visual cognitive development. In *Trends in Cognitive Sciences*, **9**, 144-151.

Mumford, D. and Shah, J. (1985). Boundary detection by minimizing functionals. In *IEEE Conf. on Computer Vision and Pattern Recognition*, San Francisco.

Needham, A. (1998). Infants' use of featural information in the segreation of stationary objects. *Infant Behavior and Development*, **21**, 47-76.

Needham, A. (1999). The Role of Shape in 4-Month-Old Infants' Segregation of Adjacent Objects. *Infant Behavior and Development*, **22**, 161-178.

Needham, A. (2001). Object recognition and object segregation in 4.5-month-old infants. *Journal of Experimental Child Psychology*, **78**:3–24.

Ostrovsky, Y., Andalman, A. and Sinha, P. (2006). Vision following extended congenital blindness. *Psychological Science*, **17**, 1009-1014.

Pratt, J. C., and Richards, R. D. (1968). Bilateral secondary congenital aphakia. *Arch. Ophthalmol.* **80**(4):420-422.

Shi, J. and Malik, J. (1997). Normalized cuts and image segmentation, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 731--737, San Juan.

Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: Natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences USA*, **98**, 1935-1940.

Sinha, P., Balas, B. J., Ostrovsky, Y. and Wulff, J. (In press). Visual object discovery. In *Object Categorization: Computer and Human Vision Perspectives*, Dickinson et al. (Eds), Cambridge University Press.

Tu, Z., Chen, X., Yuille, A.L., and Zhu, S.C. (2003). Image Parsing: Unifying Segmentation, Detection, and Recognition. *Proceedings of the IEEE Intl. Conf. on Computer Vision.*

Ullman, S. (1996). *High-level vision.* MIT Press: Cambridge, MA.

Valvo, A. (1971). Sight Restoration after Long-Term Blindness: The Problems and Behavior Patterns of Visual Rehabilitation. American Foundation for the Blind.

Von Senden, M. (1932). *Space and Sight: The Perception of Space and Shape in the Congenitally Blind Before and After Operation.* Reprint, Glencoe, IL: Free Press, 1960.

Wertheimer, M. (1938). Laws of organization in perceptual forms (partial translation). In *A Sourcebook of Gestalt Psychology*, W. Ellis (Ed.). Harcourt Brace and Company, pp. 71–88.

Chapter 4. An argument for Dynamic Bootstrapping as a theory of visual development

Combining the insights we gain from our late-onset vision subjects and the evidence from existing infant studies, we outline here a candidate theory of the development of object-level visual segmentation.

4.1 Introduction

When a human adult with normal vision looks upon Figure 4.1, three novel objects upon a textured background are readily apparent. The visual system is able to perceive the boundary between each object and its background based on a variety of cues, including changes in color, texture and luminance, as well as the continuity of the subtle borders created by these changes. It is important to notice that these are novel objects, thus we must rely largely on general, “low-level” cues to accomplish the task of segmenting these novel objects from their novel background, rather than higher level knowledge gained from experience with these particular objects.

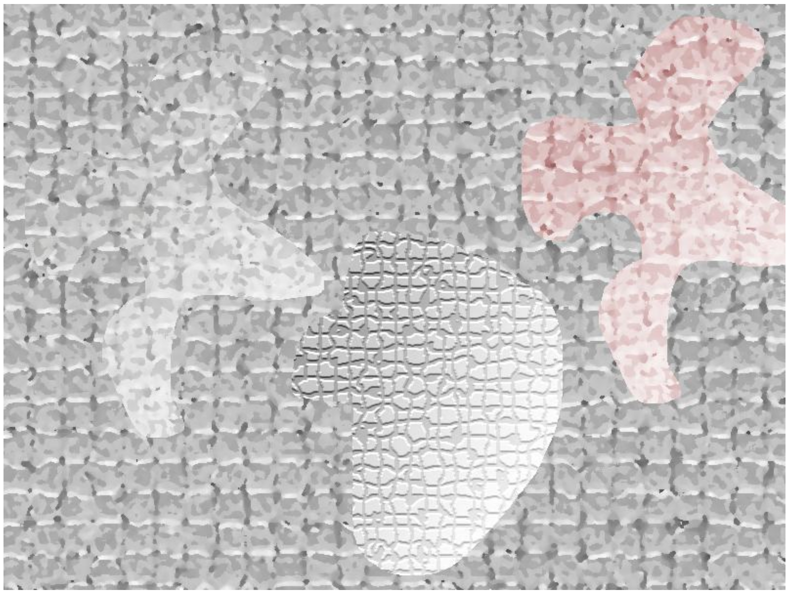


Figure 4.1 A static "scene" that is easily parsed by a casual observer into three novel objects and a textured background. Cues that may enable this parsing include differences in luminance, color and texture as well as the continuity of edges formed by these cues. Cues such as these, commonly referred to as “Gestalt” laws or rules, seem important for the decomposition of the visual world into relevant parts, objects, and backgrounds. Are the processes which enable this parsing endowed through evolution, or are they sculpted through experience with the natural world?

The ease with which a normal observer can accomplish this task stands in marked contrast to the difficulties which newly sighted individuals (both infants and late-onset vision patients) encounter, as discussed in previous chapters. Clearly, at least some of these abilities develop over time. In this chapter we argue that this progression of the visual system is not simple maturation of innate visual skills, but rather a consequence of learning from the visual environment.

4.1.1 The Historical Perspective

No serious researcher would disagree that the “visual brain” is plastic and that the representations of objects, faces, scenes and textures are shaped to some degree through visual experience. The literature on perceptual learning is vast. A few prominent examples are shown in Figure 4.2. These findings are an example of long term or permanent changes to the visual system brought about by repeated exposure (or, in some cases, just one exposure) to these displays. An even more obvious example of this phenomenon might be the quotidian act of object or face recognition – the more one is exposed to an individual’s face, for instance, the quicker and more reliably that individual can be identified.

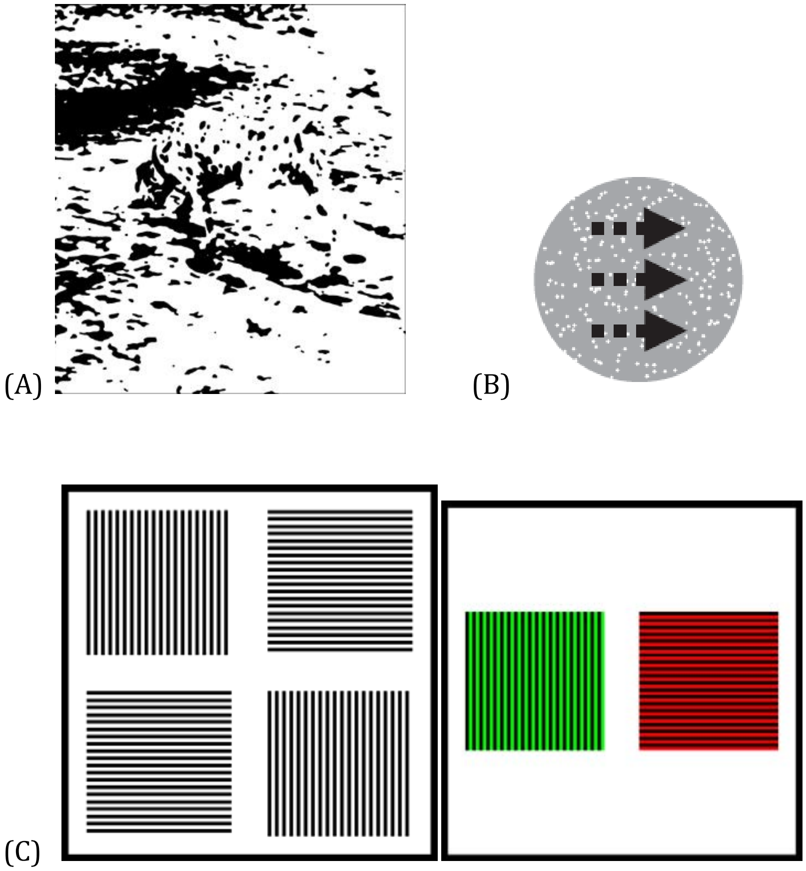


Figure 4.2 (A) Dalmatian dog illusion (Gregory, 1970) (photo taken from (Rubin, Nakayama, & Shapley, 1997)). Although perceptual organization of this scene is very difficult upon first sight of this image, the embedded dog is instantly recognized on subsequent viewings once the dog is identified to the observer just once. (B) Perceptual learning accounts for increased performance for tasks such as detection of changes in motion ((Ball & Sekuler, 1987); this image and review of similar tasks of perceptual learning in (Fine & Jacobs, 2002)). (C) The McCollough effect. After only a short viewing of the colored oriented bars on the right, a color after-effect is observed in the squares on the left, even after 2,040 hours between initial presentation and test (Jones & Holding, 1975).

Such learning is not particularly controversial. Our primary concern here, however, is not the representation of object instances (which are clearly experience dependent) or the ability to become more skilled at a particular task, but rather the fundamental development of the "rules" or "heuristics" which parse any given scene – novel or familiar – into its constituents. These rules can be thought of as the elements of perceptual organization. The mere existence of perceptual learning suggests that it may be possible for perceptual organization principles to be learned or at least modified, but this question has raised considerable controversy throughout the history of vision science.

4.1.1.1 The Gestaltists

The Gestalt "laws" of perceptual organization (Koffka, 1935; Wertheimer, 1923) identify a number of important grouping principles that have scaffolded modern perceptual theory. These principles, illustrated in Figure 4.3, include *proximity* (A), *similarity of color* (B), *similarity of size* (C), *common fate* (D), *good continuation* (E), and *closure* (F) (Palmer, 1999). Given processes to compute heuristics such as these, a visual scene can theoretically be decomposed into objects or object sets. The goal of the Gestaltist philosophy is to identify the complete set of such principles which underlie all of visual perception.

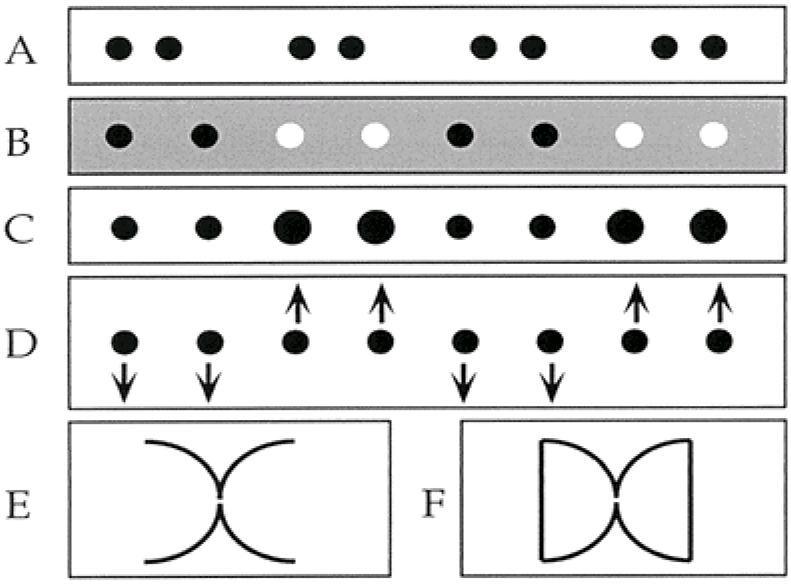


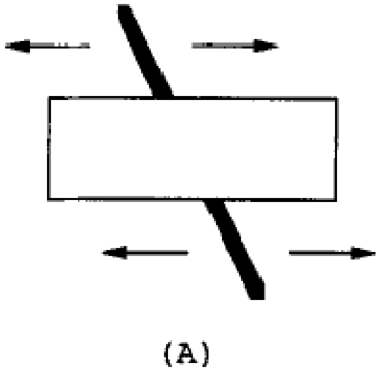
Figure 4.3 The Gestalt principles of perceptual grouping. Image from (Palmer, 1999).

The Gestaltist view seems to suggest a nativist position regarding the provenance of these perceptual building blocks; that is, the infant brain is born with these capabilities, and it is the job of the vision scientist to merely discover the processes endowed by evolution. This is in contrast to the empiricist position, wherein experience with the visual world guides the formation of these processes. Gordan and Slater (1999) note that Wolfgang Köhler, one of the founders of Gestalt psychology in the earliest decades of the 20th century, took the extreme nativist view. The debate hardly abated several decades later when Kanizsa (1979) wrote that "neither of the two theses [nativist versus empiricist] has clearly prevailed over the other, and the situation is still far from clear" (p. 25). His bias, however, seems toward the nativist view, as he outlines a series of phenomena and diplomatically concludes that “the results of these experimental situations do not seem to be compatible with an extreme empiricist explanation of perceptual organization” (p. 27).

4.1.1.2 Recent developments

Although the Gestaltists made great strides in understanding the underlying “rules” of visual perceptual organization, it was the subsequent studies with infants which arguably provided the first significant empirical evidence attempting to resolve the nativist versus empiricist question in visual perception. Research with normal adults can serve to identify the basic properties of visual perception but not whether the processes which identify those properties are innate or acquired through experience. Uncovering what visual abilities are present at or near birth, however, can give us much greater insight into the debate.

One of the earliest and most important set of experiments in this domain was performed by Kellman and Spelke (1983). These researchers found that common fate motion, one of the principles in the Gestalt framework, was highly effective at grouping the visually segregated parts of an occluded object in 4-month old infants (Figure 4.4). Later studies found similar results in even younger infants (2 month olds) (Johnson & Aslin, 1995). The very young age at which this ability presented itself suggested an innate ability, and Spelke (1990) hypothesized that infants make use of innate processes to create object models based on motion and surface segmentation properties; these models, in turn, provide the basis for learning other Gestalt grouping cues, such as symmetry and regularity—cues which infants did not reliably employ.



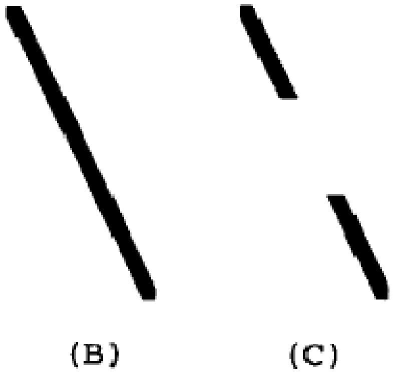


Figure 4.4 Infants observe a "rod" occluded by a box in its center (A). When the rod undergoes translational motion while "behind" the box, infants perceived the rod as unbroken, as tested in a novelty looking preference test (B versus C) following habituation. Without motion, 2-4 month old infants perceive the rod as two individual pieces. Image from (Johnson & Aslin, 1995).

Subsequent studies, however, suggested a more complex interpretation. The ability to perceive object unity in displays such as Figure 4.4(A) *without* motion did not develop until 6.5 months of age and the ability to interpolate the missing portion of the rod correctly was not evident until 8 months (Craton, 1996). Were these abilities merely slower to *mature* in the infant brain than the ability to process common fate motion (but still innate)? Or does the visual system *learn* these abilities through experience with the visual world? It is difficult to discern between these two possibilities by studying normal infants, as the infant brain is maturing and learning simultaneously. Hence, we turn our attention to the findings of our work in Project Prakash.

4.2 Motivation towards an empiricist theory of visual development

Several findings described in earlier chapters have important implications for any theory of the development of visual parsing.

Firstly, the case study of subject SRD shows that after 12 years of visual deprivation from birth, functional vision is largely undeveloped, yet after 20 years of visual experience, visual function becomes almost fully developed. A pure nativist theory of visual development would predict either (1) a critical period of maturation, after which vision would no longer develop under any circumstances, or (2) the near-immediate onset of visual function after recovery of optical function, due to the maturation of visual function along with other neurological function. Neither of these predictions matches the observed behavior of SRD. An alternative nativist explanation, in light of these facts, is that visual function is innate and comes about through brain maturation *normally*, yet the brain, being resilient, *can* acquire visual function through experience by utilizing alternative strategies. Although in principle this is a possible explanation, it seems unusual and wasteful that an ability which can demonstrably be learned (and learned quite well) is instead given a privileged innate module, which itself takes at least 8 months to develop in the normal infant, based on the evidence from the previous section.

Secondly, our studies with Project Prakash patients immediately following sight onset demonstrate the lack of processes capable of perceptual organization in any but the most uncluttered, pre-segmented scenes, with the notable and important exception of common motion grouping. Moreover, our longitudinal studies of these same subjects reveal that more advanced processes are indeed acquired after visual experience (with significant increase in function after approximately one year). This pattern matches the ordinal time course of infants—common fate motion grouping followed by grouping via static surface features, yielding some evidence against the notion that experience after a "critical period" leads to a parallel and different visual system.

Finally, studies with adolescents and adults, as opposed to infants, allow us to discount the influence of cognitive knowledge about the physics of the world, which feature in some theories of visual development (e.g., Spelke (1990)). Our subjects have fully developed knowledge about their physical world through their haptic experiences, as well as their formal education, yet this knowledge does not significantly inform their visual interpretations.

These findings motivate a theory of visual development which focuses on the extraction of Gestalt-like heuristics via learning based on visual experience.

4.3 The Dynamic Bootstrapping hypothesis: The development of visual perceptual organization abilities through dynamic information

Our findings suggest a general framework to describe how a newly-sighted individual progresses from an overly fragmented world of bits and pieces to a world of cohesive objects. The foundation of our theory is the somewhat counter-intuitive notion that dynamic information bootstraps static perceptual organization. In other words, the information that a visual system receives from objects in motion "teach" the cues that can be used to process static scenes.

The framework is as follows:

1. The principle of common fate motion as a grouping cue is learned very quickly through simple, local Hebbian learning or a similar process. This learning may be driven by the presence of simultaneous change in fragments belonging to objects that are moving, to which the brain is particularly sensitive. Alternatively, we may take the common fate motion principle as one of the few innate processes without fundamentally altering the rest of the framework.
2. Object fragments are grouped together early in development into "proto-objects" via common fate motion cues. These proto-objects serve as the "ground truth" for the visual system's notion of objectness.
3. Given a "signal" that certain parts should bind together in the form of these proto-objects (and also a negative signal that certain parts do not bind together, when object parts belong to two separate proto-objects, for instance), the visual system can extract static regularities (i.e., regularities in image characteristics not dependent on motion) and, critically, *associate these regularities with object cohesion.*
4. These static regularities can eventually be used to process a scene without dependence on motion cues.

This "Dynamic Bootstrapping" hypothesis aims to describe the development of Gestalt-like perceptual organization cues via learning from the visual environment with the notion that one set of learned processes can bootstrap the next, with motion processing as the foundation of the development of visual object concepts.

It should be noted that Prodöhl and colleagues (2003) have recently proposed a similar hierarchical framework for the learning of Gestalt rules, but heretofore there has been a lack of evidence as to its viability. It is our hope that the work in this thesis serves to encourage the acceptance of this model and to provide empirical direction to further research. Furthermore, although the aforementioned model characterizes the learned cues as methods for segmenting objects from the background, we emphasize here the role of defragmentation – the process of combining objects parts and surface fragments into a whole. Admittedly, these goals are often complementary, but our experience with Project Prakash patients suggests greater deficits in oversegmenting the surface of an object than in determining the boundary between object and background. For instance, when asked to identify the object(s) in a scene depicting a cow in a pasture, subject SK correctly identified the black-and-white patches of the cow, but individually as separate objects (though a few dark cast shadows also confused him). He did not identify the background grass and sky as a myriad of individual objects. These distinctions warrant further investigation and are not taken up here.

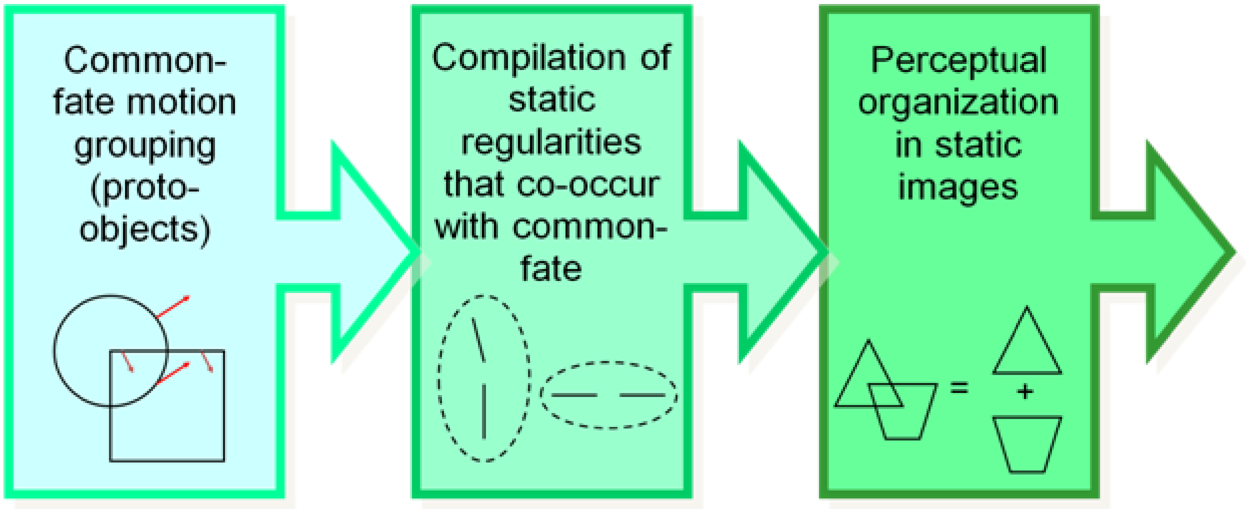


Figure 4.5 The Dynamic Bootstrapping hypothesis. A framework for how the visual system may learn perceptual organization cues by starting with just a few simple heuristics.

4.4 Evidence for the Dynamic Bootstrapping hypothesis

The Dynamic Bootstrapping hypothesis states that motion processing provides an “interpretation” of a scene which is used to learn static processes that can subsequently be used independently. Although we cannot observe this learning process directly, we examine here indirect evidence and arguments which lend credence to this hypothesis.

4.4.1 Empirical/developmental evidence

A key requirement of the hypothesis is that motion processing abilities be available early in the developmental timeline. As summarized below, this requirement seems to be met in human development.

A newborn infant, 1-3 days old, exhibits smooth pursuit eye movements in response to moving stimuli (Kremenitzer, Vaughan, Kurtzberg, & Dowling, 1979). Optokinetic nystagmus (OKN) to pattern coherence develops early, exhibiting adult-like patterns of performance at 1 month of age with no noticeable improvement over the next two months (Manny & Fern, 1990).

Global motion processing, as measured using looking times at random dot patterns, matures quite early in the infant, as compared to global form processing without motion (see Braddick, Atkinson, & Wattam-Bell, 2003 for review). Infants (and adults) show a stronger, more robust response to global motion coherence than to global form in visual event-related potentials (VERPs) (Oliver Braddick & Atkinson, 2007).Common motion is the first and dominant cue identified for object part grouping in infants, as evidenced in the long line of studies discussed in section 4.1.1.2 by Kellman, Spelke, Johnson, Aslin, and others based on the rod-and-frame tasks of Kellman and Spelke (1983). These findings primarily focus on the binding of an object partially obscured by an occluder, implying a role for motion in the 3D interpretation of a scene. Our own work, discussed earlier, with newly sighted adolescents and adults after long term, congenital visual deprivation (Ostrovsky, Meyers, Ganesh, Mathur, & Sinha, 2009) highlights an even more fundamental role of motion in binding fragments of an object that differ in texture even in the presence of a single surface. Furthermore, the ability of our subjects to use this cue at such an advanced age of sight onset confirms the resiliency of the motion cue— underscoring its importance.

One prediction of the Dynamic Bootstrapping hypothesis is that congenital deficiency in motion processing (sometimes called akinetopsia) should result in profound impairment of visual function. Such cases are exceedingly rare, but at least one such case (Ahmed & Dutton, 1996), though largely anecdotal, does suggest such a dysfunction. For instance, a staircase appears to this child to be a confusing collection of lines. Such cases provide an opportunity to test directly the *necessity* of motion processing for learning other perceptual cues. Attempts to locate more such cases have so far been unsuccessful.

4.4.2 Ecological/statistical evidence

A further requirement of the Dynamic Bootstrapping hypothesis is that the statistics of the visual world support the possibility that static grouping cues may be learned from motion cues. Although in principle this requires knowledge of the learning algorithm itself (which, of course, is unknown), it is highly suggestive if it can be shown that motion grouping cues co-occur with static grouping cues. In other words, if static grouping cues tend to occur alongside motion grouping cues (which we have already shown are efficacious early on in development), it is reasonable that some learning process would eventually learn these static cues as well.

Indeed, our work has shown that at least one important Gestalt grouping cue – contour continuation – is much more readily learnable by making use of the motion signal (Ostrovsky, Wulff, & Sinha, 2007; Sinha, Balas, Ostrovsky, & Wulff, 2009), especially over longer distances. Others have shown, with biologically-inspired neural network models, that short-range collinearity and curvilinearity can be learned from local motion signals (Prodöhl et al., 2003).

4.4.3 Logistical evidence

Finally, we offer a logistical argument for the necessity of motion in our hierarchy. In the case of contour continuity, mentioned above, common fate motion provides a more robust statistical dataset than statistics based solely on collinear co-occurrence in static scenes. However, static statistics do provide *some* information that could conceivably lead to the learning of contour continuity (Sigman, Cecchi, Gilbert, & Magnasco, 2001), and thus the *necessity* of motion information is still in question.

In light of this, we offer one example of a cue which seemingly cannot be learned in principle from the static information available to an infant or newly sighted individual. In Figure 4.6, the reader will find three similar images. Images (A) and (B) are consistent with a 3D-wireframe (Necker) cube interpretation, behind three rod-shaped occluders. Image (C), though highly similar, does not allow for such a perceptual interpretation. Now, suppose that the cube in (A) and (B) is completed due to the principle of contour continuity. If so, then the figure in (C) should be similarly completed, as the same contours of the cube exist in that image. Contour continuity on its own cannot be the appropriate explanation, at least for the figure in (B). The explanation likely involves the junction configuration at the intersection of the cube and its occluders.

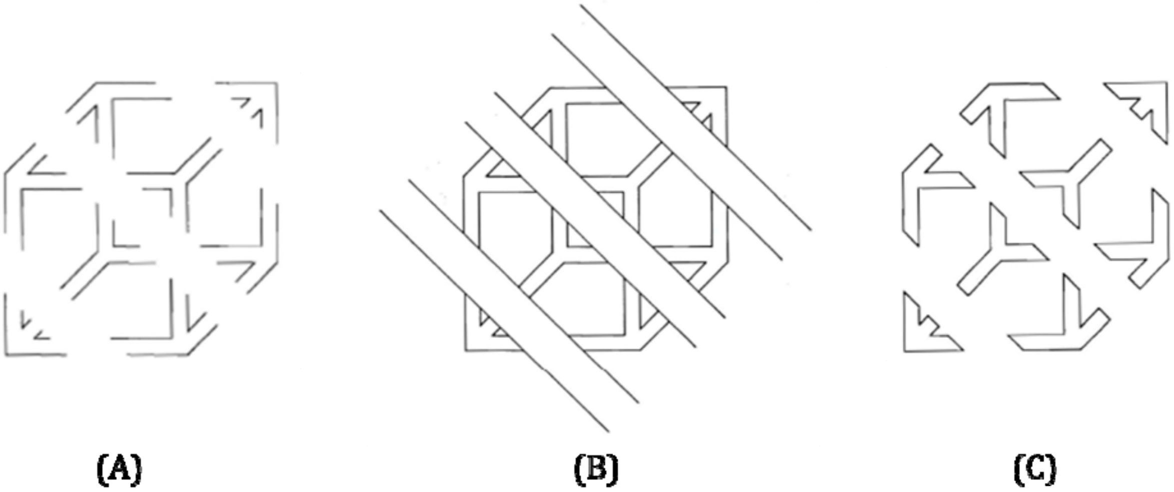


Figure 4.6 Images (A) and (B) suggest a Necker cube (3D-wireframe cube) behind three rod-shaped occluders. Contour continuation could account for this effect. However, image (C) does not suggest a Necker cube, contrary to what a contour continuation explanation would predict. See text for why motion cues seem necessary in a learning-based account of this effect. (Images from Kanizsa (1979, pp. 7-8).)

Now, the question remains: How is this junction cue learned through experience? If an object such as the cube were in motion behind the occluders in (B), the common fate motion cues would identify the appropriate contours of the cube to bind together in the presence of the co-occurring T-junctions. If the occluding bars were to move instead of the cube, a similar learning would occur, but for the surface of the bars, rather than the cube. The junctions in (C), however, would not likely occur in any kind of physical scenario caused by occlusion. Thus, motion would likely send the (appropriate) signal that these object parts should not bind together into a whole.

Without the benefit of motion, statistical learning seems nearly impossible in this scenario. The static display does not indicate to which surface the regions surrounding the junction belong. Although some sort of alignment of T-junction may be learned, this should result in the binding of all surfaces together. There is no cue to decouple the occluder from the cube. In principle, the cube may have a different texture from the occluder, allowing the pieces of the cube to bind, but evidence from infants and Project Prakash patients precludes the use of this cue in early vision. Motion seems the only usable cue to inform the learning of this set of principles.

4.5 Discussion

The mechanisms of the development of visual perception are crucial to the understanding of visual perception itself. Although vision scientists have identified many abstract “rules” of vision, a rigorous definition of how these rules are applied has so far eluded us. By studying the genesis of these rules, we gain insights into their implementation. We also learn much about what can go wrong in development, providing knowledge that might lead to more effective therapies for conditions such as amblyopia, visual perception problems stemming from late-onset vision, visual abnormalities due to stroke, or even some of the lesser known perceptual problems of seemingly unrelated neurological disorders such as Williams Syndrome (Atkinson et al., 1997) and Autism Spectrum Disorder (Spencer et al., 2000).

The debate over which aspects of visual processing are innate and which are learned is by no means answered here, but the body of evidence summarized in this chapter forwards two goals:

1. By decoupling biological maturation of brain mechanisms from experiential learning, we can emphasize with some confidence the prominent role of learning in visual development.
2. By combining evidence about (a) the developmental timeline of certain visual processes in both infants and late-onset vision patients and (b) the ecological statistics of visual information, we suggest a model which links the learning of static figural cues to (possibly innate) dynamic cues.

Several aspects of our model remain to be described. For instance, is the ability to process dynamic cues innate or itself learned via a more fundamental process? Is there a hierarchy of perceptual organization cues, where one set of cues enables the learning of the next; or does motion play a privileged role as “teacher,” providing the direct basis for learning all perceptual organization cues? Current data do not suggest any particular answer to these important questions. Furthermore, does this model apply to all of vision – including the processing of texture, spatial relationships, faces – or is it limited to object segmentation and defragmentation? Again, current evidence provides no answer.

To help clarify these open issues, it would be useful to study patients with motion processing deficits at various onsets throughout development (a rare scenario), or to debilitate motion processing in animals, either surgically or by limiting vision to static, 2-dimensional scenes. By manipulating access to the motion signal at various stages in visual development, we can determine the selective effects of motion at each stage of the conjectured hierarchy of perceptual cue acquisition.

Further progress can be made with computational models of cue learning. As described here, some simulations already suggest that contour continuity is a cue which can be learned more quickly via common motion information than without relying on motion. Modeling the learning of other cues using the motion signal, correlational statistics, and static cues determined empirically to be earlier in the hierarchy would help us determine the relative importance of each cue to the learning of the next.

Dynamic Bootstrapping is a framework meant to guide the development of more refined models of perceptual cue acquisition, raising perhaps more questions than it answers, but providing concrete directions for further inquiry in the empirical and computational domains.

4.6 References

Ahmed, M., & Dutton, G. N. (1996). Cognitive visual dysfunction in a child with cerebral damage. *Dev Med Child Neurol*, *38*(8), 736-9.

Atkinson, J., King, J., Braddick, O., Nokes, L., Anker, S., & Braddick, F. (1997). A specific deficit of dorsal stream function in Williams' syndrome. *Neuroreport*, *8*(8), 1919.

Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research*, *27*(6), 953-965. doi:10.1016/0042-6989(87)90011-3

Braddick, O., & Atkinson, J. (2007). Development of brain mechanisms for visual global processing and object segmentation. In *From Action to Cognition* (Vol. 164, pp. 151-168). Elsevier. Retrieved from http://www.sciencedirect.com/science/article/B7CV6-4PTMWX9-C/2/224f69b9f5a450aa99f01fd1c866e148

Braddick, O., Atkinson, J., & Wattam-Bell, J. (2003). Normal and anomalous development of visual motion processing: motion coherence and [`]dorsal-stream vulnerability'. *Neuropsychologia*, *41*(13), 1769-1784. doi:10.1016/S0028-3932(03)00178-7

Craton, L. G. (1996). The development of perceptual completion abilities: infants' perception of stationary, partially occluded objects. *Child Dev*, *67*(3), 890-904.

Fine, I., & Jacobs, R. A. (2002). Comparing perceptual learning tasks: A review. *Journal of Vision*, *2*(2), 190–203.

Gordan, I., & Slater, A. (1999). Nativism and empiricism: The history of two ideas. In *Perceptual Development* (p. 73). Psychology Press.

Gregory, R. L. (1970). The intelligent eye.

Johnson, S. P., & Aslin, R. N. (1995). Perception of object unity in 2-month-old infants. *Developmental Psychology. Vol. 31(5)*, *31*(5), 739-745.

Jones, P. D., & Holding, D. H. (1975). Extremely long-term persistence of the McCollough effect. *Journal* *of Experimental Psychology: Human Perception and Performance*, *1*(4), 323-327. doi:10.1037/0096-1523.1.4.323

Kanizsa, G. (1979). *Organization in vision : essays on gestalt perception*. Praeger special studies. New York: Praeger.

Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognit Psychol*, *15*(4), 483-524.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace & World.

Kremenitzer, J. P., Vaughan, H. G., Kurtzberg, D., & Dowling, K. (1979). Smooth-Pursuit Eye Movements in the Newborn Infant. *Child Development*, *50*(2), 442-448.

Manny, R. E., & Fern, K. D. (1990). Motion coherence in infants. *Vision Research*, *30*(9), 1319-1329. doi:10.1016/0042-6989(90)90006-7

Ostrovsky, Y., Meyers, E., Ganesh, S., Mathur, U., & Sinha, P. (2009). Visual Parsing After Recovery From Blindness. *Psychological Science*, *20*(12), 1484-1491. doi:10.1111/j.1467-9280.2009.02471.x

Ostrovsky, Y., Wulff, J., & Sinha, P. (2007). Learning static Gestalt laws through dynamic experience [Conference Abstract]. *J. Vis.*, *7*(9), 315.

Palmer, S. (1999). : Gestalt Perception. In *The MIT Encyclopedia of the Cognitive Sciences* (pp. 344-346). Cambridge, Massachusetts: The MIT Press. Retrieved from https://cognet.mit.edu/library/erefs/mitecs/palmer.html

Prodöhl, C., Würtz, R. P., & von der Malsburg, C. (2003). Learning the Gestalt Rule of Collinearity from Object Motion. *Neural Computation*, *15*(8), 1865-1896. doi:10.1162/08997660360675071

Rubin, N., Nakayama, K., & Shapley, R. (1997). Abrupt learning and retinal size specificity in illusory-contour perception. *Current Biology*, *7*(7), 461-467. doi:10.1016/S0960-9822(06)00217-X

Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: Natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences of the United States of*

America, *98*(4), 1935-1940. doi:VL - 98

Sinha, P., Balas, B. J., Ostrovsky, Y., & Wulff, J. (2009). Early stages of object categorization. In S. J. Dickinson, A. Leonardis, B. Schiele, & M. Tarr (Eds.), *Object categorization: Computer and human vision perspectives* (pp. 301-323). Cambridge, England: Cambridge University Press.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, *14*(1), 29-56. doi:10.1016/0364-0213(90)90025-R

Spencer, J., O'Brien, J., Riggs, K., Braddick, O., Atkinson, J., & Wattam-Bell, J. (2000). Motion processing in autism: evidence for a dorsal stream deficiency. *NeuroReport*, *11*(12), 2765.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychological Research*, *4*(1), 301–350.

Chapter 5. Conclusion

In these chapters, we have documented several cases of visual recovery after congenital blindness, with testing spanning from mere days to 20 years after treatment. Our prodigious battery of experiments and several stages of longitudinal follow-up constitute the most extensive study of its kind known to date. We have shown that it is possible, even after 12 years of visual deprivation, to have a near-normal recovery of high-level visual function, and we have confidence that this same level of achievement can be extended to our subject with 29 years of deprivation, owing to his continued improvement paralleling that of our younger subjects (albeit somewhat slower). Furthermore, we have discovered a few of the incremental stages through which visual development must pass, with insights into the repercussions to real-world object parsing. Finally, we compiled our discoveries with those from infant research to propose a theory of visual development based on learning from a bootstrapping process which creates proto-object representations from common motion binding.

We conclude in this chapter with the implications of our findings for the notion of a “critical period” in visual development and the nature versus nurture debate. Lastly, we discuss the future directions inspired by this work.

5.1 Thoughts on the critical period

The notion of a critical period for visual development—a special age after which the visual system can no longer develop useful sight – has been the dominant thinking for some time. Based on our experiences in Project Prakash, clinicians in India, who come across many cases of childhood blindness, will recommend against treatment for children who were born blind and remained untreated past the age of 5 or 6. This thinking, which is also in line with rich-country practice, stems from the historical literature on late-life eye surgeries, which often report the inability to attain normal adult-level sight, and perhaps from their own few experiences of late-life treatment.

Yet, this basis for clinical practice is misleading. In many cases in the literature, follow-up was inadequate. As we have seen, the development of vision – even in the normal infant! – can take many months or well over a year for an adult. Clinicians in India who have treated such cases likely did not have the time or opportunity to check on the progress of their patients. They would have seen a child who, after treatment, had little ability to use his vision, and likely little progress would have been made on the few visits afterwards. If the child did acquire vision, there would be little incentive to return to the clinic or hospital[1]. The common belief that all of visual development is a product of brain maturation led most clinicians and many researchers to dismiss the possibility that recovery could take a significant amount of time.

Another potentially misleading complication in the literature pertains to subject selection. Some subjects receiving late-life treatment for blindness such as MM (Fine et al., 2003) and a 34 year old Zambian man (Carlson, Hyvärinen, & Raninen, 1986) show little progress after treatment. Both subjects became blind after the age of 3, and thus are not cases of congenital blindness. Carlson and colleagues attempted their rehabilitation regimen, which was successful in the case of a 23 year old woman who lost her vision at less than a year old (Carlson & Hyvärinen, 1983), on the Zambian man but achieved no success in this case. They blamed the lack of progress on his apparent lack of motivation[2], but MM (a co-author on the research report about his own case) by all accounts was highly motivated in his recovery. Moreover, lacking a control group, there is no evidence that the 23 year old woman benefited directly from the rehabilitation regimen[3], so rehabilitation was unlikely to be a factor.

Paradoxically, subjects who had had significant early exposure to vision but were blinded later in childhood fared *poorer* than the congenitally blind subjects in our work and that of others. Although the number of cases is few, this pattern seems consistent. How is it that a child blinded at the age of 3, and who at that age would have had a virtually complete visual system, would come to not only lose this ability as an adult but would be at a disadvantage compared to the adult who had no vision as a child? We speculate here on a possible explanation which may have its roots in a little-understood result from prosopagnosia.

---

[1] One of the notable exceptions to this was our subject SRD, whose continued visits to the dedicated eye doctors at Raghudeep Eye Clinic in Ahmedabad, India allowed for long-term monitoring of her progress. In her many years of visits, however, her case was never catalogued or published before now.

[2] Regrettably, the lack of progress in many late-life vision patients is often attributed to factors such as low motivation, stubbornness and even poor intelligence. This unfortunate attitude may likely be due to the prevailing wisdom that, if the individual's acuity is adequate, vision should simply "happen." We hope our work sheds light on why the patient has little control over their own progress. The learning that is required to make sense of the new information coming in through the eyes is, for the most part, cognitively impenetrable (i.e., beyond the level of conscious, cognitive control).

[3] The authors cite as evidence of rehabilitative efficacy the fact that the subject had no recovery of functional vision up to 9 months after treatment and that recovery began only during the 1.5 years of rehabilitation following this time. We saw a similar pattern in our subjects, however, who partook in no rehabilitation regimen.

Prosopagnosia (Bodamer, 1947) is a disorder characterized by a pronounced deficit in recognizing faces. Oddly, its patients can perform within-category object recognition, even when matched for difficulty (Farah, Levinson, & Klein, 1995). Why, then, can prosopagnosics not enroll their object-specific recognition faculties to largely overcome their face-specific deficits? One possibility is that the use of a face-specific faculty is “mandatory”. This is borne out in the “inverted face inversion effect” (Farah, Wilson, Drain, & Tanaka, 1995), where prosopagnosic subjects show an improvement in performance for upside-down faces compared to their right-side up counterparts, the inverse of the usual effect in normal subjects. It is thought that inverting a face disrupts normal facial processing, and in the case of prosopagnosics, this may lead to relatively better performance by enabling the usage of object-specific strategies which are intact.

This phenomenon may also explain the apparent inability of patients such as MM and others to recover their previously available sight. Given many years of blindness, the substrate of cortex responsible for their high-level visual processes may have lost their function or been appropriated for other functions. If, however, the pathway to this substrate remained robust (perhaps due to spurious neuronal firing or to a consequence of anatomical connectivity), the usage of this “broken” substrate may be mandatory. Thus, as prosopagnosics are constrained to process faces as faces, not as objects, these late-vision patients must make use of a broken high-level visual system. Congenitally blind patients, however, may not have formed these robust pathways, and hence their newly forming visual system in adulthood may form elsewhere and become accessible via new pathways.

Regardless of whether this theory of mandatory usage of a broken high-level vision substrate is correct, the distinction in the clinical outcome of late-life treatments for the congenitally blind versus individuals blinded in early childhood is an important one. The ramifications of this distinction are more than academic. Adult cases of sight (non-)recovery often lead to disappointment and depression, whereas our experiences with children, adolescents, and adults have been very positive, given adequate patience. This issue requires more direct study, but our tentative consideration is that cases of congenital blindness with good eye-health outcomes are appropriate candidates for treatment even into adulthood. On the other hand, cases of early-childhood blindness with long-term visual deprivation should be considered with caution. A sufficiently brief period of deprivation in these patients would certainly allow for recovery, but we do not yet have any data points to determine this maximum duration.

5.2 Nature versus nurture

We have seen that the development of many low-level visual processes is largely driven by a process of maturation endowed to humans and other animals by evolution. Anatomical and physiological evidence shows that the timeline of this development does not depend strongly on visual input, although degenerate conditions of visual stimulation can still affect vision outcomes. Behavioral evidence in long-term deprivation studies seems to indicate that the foundations of vision are present immediately in a mature subject, although not at optimum levels and assuming enough light stimulation to prevent physical atrophy.

A different picture has emerged, however, for high-level vision. After sight onset following congenital deprivation, high-level vision takes a significant amount of time to develop. In contrast to low-level vision, this developmental process may not conform to a set maturational timeline. In the case of sight onset occurring in adulthood, visual development does not occur instantly, but rather progresses in stages as does infant visual development. There are broadly four theories to account for our observations of the development of high-level vision, and we consider each in turn:

1. High-level vision is a product of maturation, endowed completely by evolution, and unfolding unconditionally according to a set timeline from birth.

   This viewpoint is untenable. If maturation were solely responsible for high-level vision, then visually deprived subjects would either have complete vision soon after recovery (due to the conclusion of the maturational process) or would never acquire vision, having passed the critical period requiring visual stimulation. Neither of these scenarios fits the evidence.

2. High-level vision develops from visual experience.

   This is our preferred viewpoint. General perceptual learning occurs throughout life, and the cortex has been shown to be highly plastic. The timeline of acquisition, and its flexibility, is consistent with a learning explanation. However, the lack of high-level vision acquisition in non-congenital cases such as MM argues somewhat against this viewpoint. Further work needs to be done in order to determine whether our theory of pre-deprivation vision causing later-life complications (discussed above) is an adequate explanation for this inconsistency.

3. High-level vision is a product of maturation, but maturation can be “paused” pending visual stimulation.

Although this theory is difficult to rule out, there is no evidence of which this author is aware to support the notion that a maturational program can remain latent for a minimum of 29 years (as in the case of SK). By contrast, a learning account such as the one mentioned above does not depend on unknown anatomical processes.

4. High-level vision matures normally even in visually deprived individuals, but years of disuse can mask its functionality. Visual stimulation allows the visual system to re-appropriate this previously unused substrate.

   This theory seems unlikely given cases such as MM. The usage which MM had of the high-level visual "module" during his first 3.5 years, under this theory, would have made it more likely that he would have recovered normal vision compared to congenital blindness cases, which is contrary to reports.

It is important to note, however, that our leaning towards a learning account of high-level vision is not meant to completely exclude *any* role of innate biases or processes; rather, the evidence points to learning as the mechanism which enables the visual system to progress from a rudimentary starting point to a highly functional ability. As a case in point, consider the recent studies of Sugita (2008), which found that infant monkeys deprived of facial stimuli from birth (but that nonetheless had normal visual experience) exhibited a high-level of discrimination abilities for both featural and configural differences in human and monkey facial stimuli. This may indicate that monkeys, at least, may have some innate facial recognition abilities, though this by no means would extrapolate to the object parsing abilities on which we focus our attention. It is also questionable whether this result applies to humans, as configural processing in particular is known to develop over many years in children (see discussion in Chapter 1), while Sugita found that it seems to peak at just 6 months of age in monkeys, with no discernible ramp-up period, at least in the ages tested. It is also difficult to determine whether the pattern of results witnessed by Sugita is "face processing" *per se* or an instance of general object recognition. These studies do show a pattern of perceptual narrowing (Pascalis, de Haan, & Nelson, 2002), akin to the "other race" effect (Cross, Cross, & Daly, 1971), to exposed human or monkey faces, dampening observed discriminability in the non-exposed category. This narrowing would indeed point to face-specific processes, but this effect was observed after one month of exposure to the respective face category (i.e., following visual experience with the appropriate category). Furthermore, this period of exposure was equally effective after 6, 12 and 24 months of facial deprivation, indicating perhaps a learning process somewhat independent of maturation (though it is also consistent with a delayed-maturation account). Regardless of whether some aspects of face processing may be innate, the development of these processes to maturity seems to involve a categorical shift after visual experience in individuals visually-deprived from birth. In face detection tasks, for instance, performance after little visual experience exhibits a pattern consistent with the performance of an image-matching algorithm, with false detections correlating with simple image-level similarity, while experience of 6-18 months yields a clearly categorical effect consistent with a more sophisticated parsing of face-specific representations (Meng, Cherian, & Sinha, 2009). Thus, visual experience may play a substantial role in the development of high-level vision processes, even if the rudiments of these processes are present soon after birth.

Even if our studies do not resolve the larger nature vs. nurture debate in the domain of vision, our work chronicles the stages of high-level visual acquisition in greater detail than has been seen before. Although some stages of this acquisition have been observed in infants in the ground-breaking work of Spelke and, later, Johnson and others (see introduction), the opportunity to take away the variables of general intelligence development and real-world knowledge has shed light on the purely perceptual aspects of vision. These insights have led us to propose a theory to bootstrap high-level visual learning based on the observed timeline in real-world acquisition.

Our conclusions are still tempered by the caveat that visual acquisition in an older child may not be a wholly accurate representation of normal infant visual development. Hopefully, our observations here can serve as the groundwork for further work in infants and animals in order to confirm (or disprove) the general applicability of our findings.

5.3 Future directions

Our work on visual development lays only the foundation for a far more ambitious research program. If our conclusion that high-level visual development is a consequence of learning from the visual environment is true, a number of interesting research questions beg to be addressed. Here, we list a few of these directions.

1. Identifying mechanisms for learning

One key aspect of the learning process is the mechanism for binding one statistical regularity (e.g., continuous contours) to a known interpretation for another statistical regularity (e.g., the “objectness” of elements undergoing common motion). One principle which is worth exploring is temporal synchronicity, or simultaneity. Synchronicity seems instrumental in very rapidly altering proprioceptive mappings via touch and vision, such as with the rubber-hand illusion (Botvinick & Cohen, 1998) where synchronous haptic and visual input "remap" the perceived positioning of a subject's hand onto an "alien" limb. Perhaps moments of sudden change in the visual domain serve as binding signals between elements undergoing change at the same time. Some of our work has pointed out the importance of such sudden "accelerations" for object binding (Ostrovsky, Leonova, & Sinha, 2008). Our work emphasizes the importance of motion information in learning, but signals for statistical learning in the static world are undoubtedly important as well.

2. Discovering the timeline of other visual perceptual abilities

Our work in Project Prakash has identified the developmental progress of a few Gestalt-like perceptual cues, as has the infant literature. As efforts proceed toward identifying a larger number of subjects and increasing their accessibility to the researchers, we will have the opportunity to probe the development of these and other cues on a finer timescale. The order in which the ability to utilize each of these cues develops would be useful for hypothesizing the dependence of one set of cues on another, though the order alone is not sufficient to prove this dependence. Experimental manipulations during within-session learning, if this is possible, may shed light on the question of causal dependence.

3. Possibilities for manipulation of the course of visual learning

If indeed many perceptual abilities are learned from the environment, it should in principle be possible to manipulate the acquisition of these abilities. This would be most useful for the late-onset vision patients, whose progress in acquiring visual function initially is somewhat slow. Behavioral interventions may speed up the rate of functional acquisition by, perhaps, appropriately concentrating training on important defragmentation cues. This kind of intervention would have applicability to our patients in the developing world, but may also become important more generally as new technologies such as retinal implants (Rizzo et al., 2001) become common place.

4. Expanding the scope of learning to other modalities

In addition to object segmentation and defragmentation functions discussed here, the newly sighted individual must bind visual representations to haptic ones. William Molyneux (1688), in a letter to the philosopher John Locke, first expressed the problem: If an individual, blind from birth, were suddenly made to see, would that individual be able to distinguish a cube from a sphere (having touched both previously) by the use of sight alone?

As with the development of perceptual organization, this cross-modal recognition ability may also be either innate or learned. Prior work has pointed to the high level of plasticity in the cross-modal domain (Held, 1965; Held & Bossom, 1961), making this a rich domain for the exploration of learned mechanisms.

We have probed the Molyneux question in our Project Prakash subjects, and the answer seems to be that the ability is not present immediately but is rapidly acquired over the course of weeks (in preparation). The mechanisms of this acquisition may shed light on how purely visual skills are acquired and *vice versa*.

5.4 Postscript

Since the work published in these chapters, we have had several more subjects, all showing similar signs of visual learning. BK is one such example. BK and his two brothers all suffered from congenital bilateral cataracts leading to a life of blindness. Years earlier, BK's older brother had received treatment for his cataracts but, sadly, passed away after complications from infection at a less competent clinic than our partner hospital, making his parents understandably reticent to pursue treatment for their remaining two sons when we presented that opportunity. Nonetheless, eventually they decided to allow BK to go forth with the normally routine operation (phacoemulsification with an intraocular lens implant).

Prior to his subsequent cataract surgery at the age of 16, BK was quiet, with seemingly poor self-confidence. Soon after surgery, his demeanor did not much change, nor did his vision. Although his acuity had improved, his usage of that vision was minimal. Several months later, his performance remained the same. On one of our most recent visits, however, BK was a changed man. Suddenly, his performance on our tests was near perfect. He beamed a smile while telling stories of his shopping for food, by sight, on his own. His posture was upright, rather than sullen, and he seemed to have a new-found confidence. A few months after BK's treatment, his younger brother received an operation as well. Now BK good-naturedly pokes fun at his brother's inability to make sense of the shapes with which he himself had had so much trouble only months earlier. His brother, happily, seems well on his way to mimic BK's progress.

These cases have changed the policies at Shroff Charity Eye Hospital (Delhi, India), where we performed most of our tests. Whereas previously the doctors were hesitant to operate on patients past early childhood, they now routinely treat patients like BK, providing them with greater opportunities in a socioeconomic reality where blindness is a far greater challenge than in the world of anyone reading this treatise. The opportunity to directly affect the lives of these individuals has been the largest benefit of my work with Project Prakash.

It has been a privilege to collaborate with the tireless doctors at the various hospitals and clinics we have visited, most of which are charity endeavors. Their daily efforts help countless children (and adults) with limited prospects in life. It is a hope that our work can have a fraction of the impact that these organizations have had over many decades. My gratitude also goes out to our many subjects, who endured much boredom but also shared in many laughs during our incessant experimentation. Sadly, one of our subjects, SRD, had her life cut tragically short in a bus accident not long after we had met her. She had endured much in her personal life, yet her spirit was courageous, kind, and giving. She will never know the impact she has had outside of her world.

5.5 References

Bodamer, J. (1947). Die prosop-agnosie. *European Archives of Psychiatry and Clinical Neuroscience*, *179*(1), 6–53.

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature*, *391*(6669), 756. doi:10.1038/35784

Carlson, S., & Hyvärinen, L. (1983). Visual rehabilitation after long lasting early blindness. *Acta Ophthalmologica*, *61*(4), 701-713. doi:10.1111/j.1755-3768.1983.tb04361.x

Carlson, S., Hyvärinen, L., & Raninen, A. (1986). Persistent behavioural blindness after early visual deprivation and active visual rehabilitation: a case report. *British Journal of Ophthalmology*, *70:*, 607 -611. doi:10.1136/bjo.70.8.607

Cross, J. F., Cross, J., & Daly, J. (1971). Sex, race, age, and beauty as factors in recognition of faces. *Perception and Psychophysics*, *10*(6), 393–396.

Farah, M. J., Levinson, K., & Klein, K. (1995). Face perception and within-category discrimination in prosopagnosia. *Neuropsychologia*, *33*(6), 661-674. doi:10.1016/0028-3932(95)00002-K

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. R. (1995). The inverted face inversion effect in prosopagnosia: Evidence for mandatory, face-specific perceptual mechanisms. *Vision Research*, *35*(14), 2089-2093. doi:10.1016/0042-6989(94)00273-O

Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., Wandell, B. A., et al. (2003). Long-term deprivation affects visual perception and cortex. *Nat Neurosci*, *6*(9), 915-916. doi:10.1038/nn1102

Held, R. (1965). Plasticity in sensory-motor systems. *Scientific American*, *213*(5), 84–94.

Held, R., & Bossom, J. (1961). Neonatal deprivation and adult rearrangement: Complementary techniques for analyzing plastic sensory-motor coordinations. *Journal of Comparative and Physiological Psychology*, *54*(1), 33-37. doi:10.1037/h0046207

Meng, M., Cherian, T., & Sinha, P. (2009). Neural basis of contextual modulation on categorical face perception. *Journal of Vision*, *9*(8), 456. doi:10.1167/9.8.456

Molyneux, W. (1688). Letter to John Locke, 7 July, in The Correspondence of John Locke (9 vols.), E.S. de Beer (ed.), Oxford: Clarendon Press, 1978, vol. 3, no. 1064.

Ostrovsky, Y., Leonova, A., & Sinha, P. (2008). Binding the pieces: Efficacies of grouping cues [Conference Abstract]. *Journal of Vision*, *8*(6), 70. doi:10.1167/8.6.70

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is Face Processing Species-Specific During the First Year of Life? *Science*, *296*(5571), 1321-1323. doi:10.1126/science.1070223

Rizzo, J. F., Wyatt, J., Humayun, M., de Juan, E., Liu, W., Chow, A., Eckmiller, R., et al. (2001). Retinal prosthesis An encouraging first decade with major challenges ahead. *Ophthalmology*, *108*(1), 13–14.

Sugita, Y. (2008). Face perception in monkeys reared with no exposure to faces. *Proceedings of the National Academy of Sciences*, *105*(1), 394-398. doi:10.1073/pnas.0706079105

5.6 Acknowledgements

In 2003, my advisor and mentor, Pawan Sinha, launched an initiative that has become a cornerstone of my life and career—Project Prakash. The idea was genius in its simplicity, yet far-reaching in its goals: Help the hundreds of thousands of blind children in India who have been overlooked or unable to receive treatment, and use our neuroscience background to learn something about the visual system as these children acquire sight. Upon first hearing the idea, it seemed so obvious that I thought, surely, someone must be doing it already. Yet, no one was. Therefore, we must.

By that point, I had seen the impact of poverty in India and other developing nations in my travels. I had also benefited, during those same travels, from the kindness and hospitality of those who seemed to have almost nothing but their incredible love for life. I was yearning to give something back, and Project Prakash offered the opportunity to both make a difference in the world and, potentially, to make a lasting impression in the field of neuroscience.

Realizing the goals of Project Prakash turned out to be immensely more difficult than we predicted at the outset. The challenges of locating documentable cases with just the right conditions while working within the constraints of our patients' responsibilities to their families, their often difficult personal situations, and the sometimes unpredictable infrastructure of an ever-developing India have been daunting at times. Yet, along the way, the search itself has touched the lives of thousands of children and adults. I will never forget the warm welcome from the many children in the villages and vision camps we visited who would storm our bus in order to have us take their picture or merely wave at them as we left. For these experiences and for the opportunity to share in the creation of the knowledge we have gained from our work, I am sincerely grateful and indebted to my advisor.

I would also like to acknowledge the other members of my committee, each of whom has contributed to my career in various ways: Patrick Cavanagh, who first ignited my interest in vision science while I was an undergraduate taking his class in experimental methods in vision; Nancy Kanwisher, who has repeatedly provided the necessary grounding to make sure I back up all my claims with hard evidence; and the inimitable Richard M. Held, who, in addition to becoming a friend, has become a surprisingly active and frequent collaborator of mine even well into his *emeritus* years and who often punctuates a scientific discourse that might have veered off into minutiae with the remarkably effective question: “What does it all mean?” – a question which, I hope, I have begun to answer.

Finally, I would like to acknowledge my family and friends who put up with my oft-peripatetic lifestyle and supported me throughout my many endeavors throughout these years.