Thinking for Seeing: Enculturation of Visual-Referential Expertise as Demonstrated by Photo-Triggered Perceptual Reorganization of Two-Tone "Mooney" Images

Jennifer M.D. Yoon (jennifer.yoon@stanford.edu)[1]
Nathan Witthoft (witthoft@stanford.edu)[1]
Jonathan Winawer (winawer@stanford.edu)[1]
Michael C. Frank (mcfrank@stanford.edu)[1]
Edward Gibson (egibson@mit.edu)[2]
Ellen M. Markman (markman@stanford.edu)[1]
[1]Department of Psychology, Stanford University, Stanford, CA 94305 USA
[2]Department of Brain and Cognitive Sciences, MIT, Cambridge MA 02139

Abstract

Two-tones ("Mooney"-esque transformations of grayscale photographs) can be difficult to recognize. However, after viewing the photographs from which two-tones were created, adults experience rapid "perceptual reorganization," and the two-tones become immediately recognizable. In contrast to the effortless post-cue recognition seen in adults, preschool-aged children are generally unable to recognize two-tone images *even when the photograph is simultaneously available*. While simple instructional and perceptual interventions were ineffective, a cognitive intervention in which children were convinced that the photo and two-tone images were transformations of the same physical object improved children's recognition. We found a similar deficit in recognition in adults from a hunter-gatherer tribe (Pirahã) with a sparse visual symbolic culture and limited exposure to modern visual media. Photo-triggered perceptual reorganization of two-tone images may therefore be a product of prolonged enculturation, reflecting visual-referential expertise. As we gain skill in representing visual correspondences, one of the surprising consequences may be the ability to literally see things we couldn't see before.

Keywords: vision; perception; recognition; development; culture; top-down effects; perceptual reorganization; expertise; referential and symbolic understanding.

Introduction

Object and scene recognition are typically rapid and automatic (Potter & Levy, 1969). When viewing impoverished images however, recognition may be difficult. One class of such images is two-tones. Two-tones are created by adjusting and thresholding photographs to render them in two values (pure black and pure white). Some of these images, similar to the famous "Dalmatian in the snow" (Gregory, 2001), can be difficult to interpret because the foreground objects are not easily segmented from the background scenes. When briefly viewed for the first time, a two-tone image may be perceived as a disorganized smattering of black and white patches. With the right cue, such as the photo from which the two-tone was derived, viewers find the image suddenly and vividly transformed into a coherent percept (Gregory, 2001). Moreover, following reorganization, it is difficult if not impossible to see the image as the meaningless array of patches it once was. This serves as a powerful demonstration of the ability of information outside the image to drive perceptual reorganization.

Figure 1. Dalmatian in the snow from Gregory, 2001.

Kovacs and Eisenberg (2004) showed 4-5 year old children two-tone images and their corresponding photos. None of the eight children were able to verbally identify the two-tone images, even with simultaneous presentation of the corresponding photo. This phenomenon is striking in that children appear to not just be slower or less accurate; rather, when viewing the two-tone images and photos side-by-side, they fail to see what adults see automatically.

In our previous work, we replicated and extended the Kovacs & Eisenbergy (2004) finding in children. For our experiments we quantified performance with a measure of "drawing accuracy": observers were asked to mark several corresponding features in the two-tone image and the matching photograph (Yoon, Winawer, Witthoft, & Markman, 2007). They received a maximum score of 1 if they were able to correctly mark corresponding parts in the two-tone and photo images, and a minimum score of 0 if no parts corresponded. Adults drawing accuracy was nearly perfect (Figure 4, Control 2), while children's drawing accuracy was very low (Figure 4, Verbal report). The low marking scores were not due to an inability to understand the task or to relatively poor drawing skills. The same children had nearly perfect performance marking

corresponding parts in easily recognizable but non-identical image pairs (Figure 4, Control 1).

Children's striking difficulty in marking corresponding parts of two-tone and photo images was observed in two subsequent experimental conditions. In the first condition, the experimenter explicitly informed the children on every trial that the two-tone and photo images were the same ("Actually, this is a picture of the same thing as this. So there is a [name of object] in this picture and in this picture"). This manipulation did not improve performance (Figure 4, Instruction). Second, we changed the order in which showed the stimuli: we first presented the photo and only presented the two-tone afterwards with explicit instruction about the relationship between the photo and two-tone. This manipulation tested the possibility that poor performance was due to "perceptual interference", that is, that an initial incorrect interpretation of the two-tone blocked reorganization. However, this manipulation also failed to improve children's performance (Figure 4, Interference reduction).

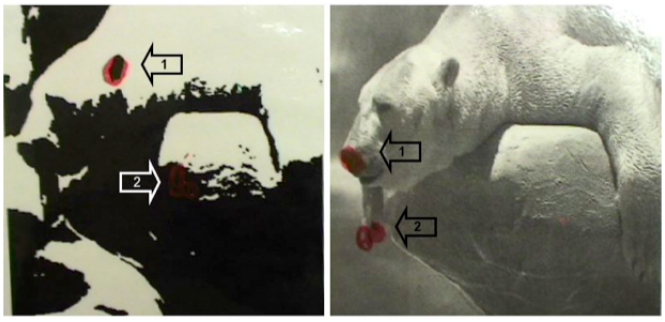

Figure 2 An example in which the child indicated a match but could not draw corresponding parts. In red, a 4-year-old has drawn the nose (arrow 1) and claws (arrow 2) of the bear in the photo, and what she claims are the same parts in the two-tone image of the bear.

One hypothesis to explain the lack of perceptual reorganization in young children is that their early visual system is not yet mature, perhaps due to a lack of fully developed long-range connections. Another hypothesis for children's deficient perceptual reorganization is that using the photograph as a cue for reorganization requires an appreciation of the dual nature of the photograph as both a concrete object in and of itself, and a representation of something else—the two-tone image. This skill is a form of visual-referential expertise entrained by a visual symbolic culture typical in the US. For example, we "read" visual symbols such as representational art, maps and models, and writing, by appreciating their relationship to the visual objects they refer to. Difficulty in such visual-referential expertise or "dual representation" (DeLoache, Miller, & Rosengren, 2007) may prevent children from capitalizing on the photo as an important cue. Here, we report two experimental results that are consistent with the referential expertise hypothesis. The first experiment shows that the likelihood of children experiencing perceptual reorganization can improve if the burden of dual representation is reduced. The second experiment shows that an adult population with minimal expertise in visual symbols has difficulty experiencing perceptual reorganization of two-tone images.

Experiment 1

The first experiment was inspired by a "shrinking machine" experiment conducted by Deloache and colleagues (1997). We showed children pairs of identical photos. One of these photos was sent through a "special machine" that led children to believe the photo *transformed* into a two-tone.

Participants Eleven preschool children (3y,9m – 5y0m, average 4y6m) from Bing Nursery School participated in the study. Four additional children were excluded for inability to finish the experiment or experimenter error.

Stimuli Two-tone images were created by thresholding blurred gray-scale photographs of a kitten, dolphins (the example image used in the Kovacs & Eisenberg, 2004), a cheetah, a dog, a polar bear, a tiger, and an athlete (Figure 3). Each trial included a two-tone image paired with the gray-scale photograph from which it was derived. Two additional practice image pairs were created using simpler image transformations. A photograph of a deer was paired with a blurred version and a house was paired with a higher contrast version. Each image was printed onto a 12 x 12 cm card.

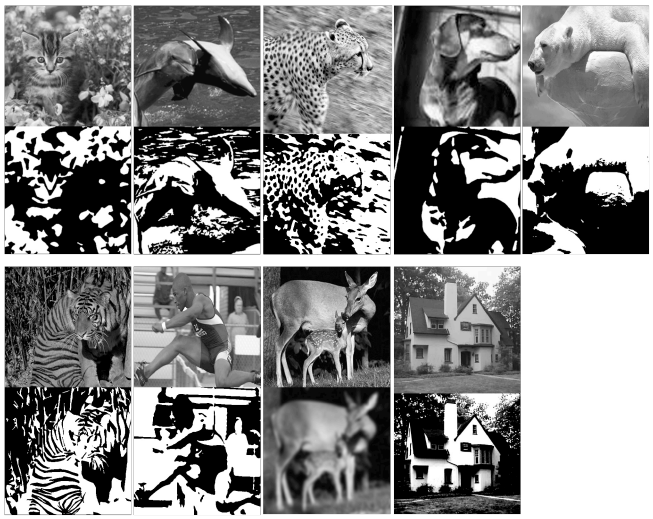

Figure 3:Full stimulus set used in Experiment 1 including seven two-tone and photo pairs, and two practice pairs (deer and blurred version, house and contrast increased version).

Procedure The experimenter sat across from the child at a small table and explained: "In this game, I'm going to show you some pictures. You just tell me what you think is in this picture. Sometimes, the pictures will be very clear and

sometimes they will look all fuzzy and blurry." Following the perceptual interference reduction manipulation (Yoon et al, 2007), we did not start trials with uncued two-tone identification. We instead began with photo identification before proceeding to cued two-tone identification and finished by asking the child to draw corresponding parts of the two-tone and photo images.

Experiments began with photo cue identification. The experimenter placed a photo card before the child and asked, "What do you think this pictures is?" She then placed an identical photo next to it and asked "What about this picture? What do you see?" The experimenter sent one of the two identical photos through a "special machine" that "makes pictures look a little different, even though they are really still the same." The machine was a modified HP Business Inkjet 2300 printer that would feed the photo card through a hidden slot behind the printer and would output a two-tone card from a hidden tray within the printer. Special care with timing and presentation gave the impression that the 12x12cm card that came out of the machine was the very same card that went into the machine

Photo cue identification was followed by cued two-tone identification. In this stage, the experimenter placed the two-tone card that emerged from the machine next to its corresponding photo. She explained, "Look what the machine did to the picture! Even though it looks different it's still a pictures of the same [name of object]." Finally, children were invited to draw corresponding parts of the two-tone and photo image.

Drawing measure The experimenter picked two or three features of the named image and asked, "Can you draw for me where the [feature] is in this picture?", pointing to the photo. Then the experimenter asked, "Can you draw for me where the same [feature] is in this picture?", pointing to the two-tone image. The experimenter gave positive feedback no matter what the child drew, "Wow, that looks great! Thank you."

For each trial, drawings on the paired two-tone and photo images were then given a single score of 0, 0.5 or 1. Drawings earned a score of 1 if all parts drawn on the two-tone image corresponded correctly to the parts drawn on the photo. Drawings earned a score of 0.5 if at least one part (e.g., nose) drawn on the two-tone image corresponded exactly with the part drawn on the two-tone image, but some remaining drawn parts did not correspond. Drawings received a score of 0 if no parts drawn on the two-tone image corresponded to parts drawn on the photo. Drawings were scored by the first author (JMDY).

Results and discussion The average drawing scores on all test trials was 0.66 (SD = 0.14) (Figure 4, bar outlined in green). This score was lower than the average drawing score on practice trials (0.97, SD = 0.08; $t(10) = 5.28$, $p < .001$, paired two-tailed t-test). However, this score was also significantly higher than the average drawing score on all test trials in a similar design that did not reduce dual representation demands (Figure 4, Interference reduction study, 0.33, SD = 0.19; $t(19) = 5.24$; $p < 0.001$, unpaired two-tailed t-test). Even though children in the interference reduction experiment were explicitly instructed about the correspondence between two-tone and photo image pairs and asked to identify the two-tone only after first seeing the photo, their drawing scores did not improve (Yoon et al, 2007). In contrast, removing the need to represent the correspondence between two distinct visual objects by convincing children the photo physically *transformed* into the two-tone was sufficient to improve their perception of the two-tones.

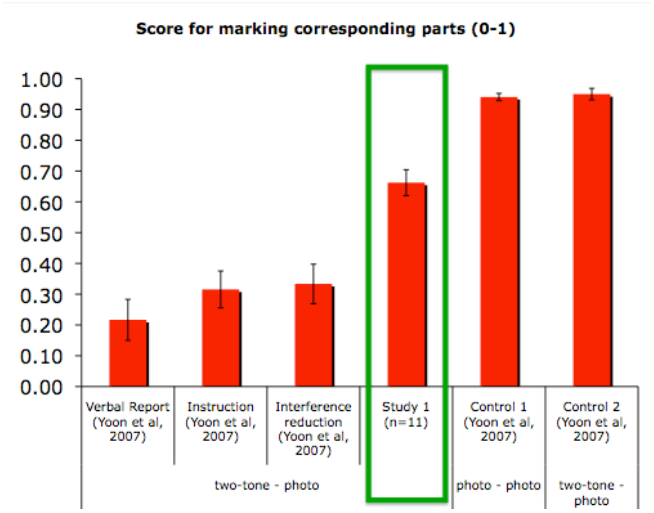

Figure 4: Performance in drawing corresponding parts on image pairs. Children are significantly worse at drawing corresponding parts of two-tone and photo image pairs compared to adults and compared to children drawing on two identical photographs. Neither instruction about the identity between the two-tone and photo images nor reducing perceptual interference helped to improve drawing scores significantly. The cognitive intervention described in Study 1 (outlined in green), however, doubled performance.

Pirahã Adults and Controls

Although an experimental manipulation improved children's performance in Experiment 1, it may still be the case that physiological maturation is sufficient to guarantee perfect performance in adults. Physiological maturation of the visual system, regardless of cultural environment, could explain differences in perception between children and adults. However, children may also become more strongly enculturated into the practices of perceptual inference and interpretation accepted in their particular community over time, similarly predicting differences in how children and adults perceive the world. But is the rapid perceptual reorganization reported by US adults a necessary consequence of having a mature visual system or is it the result of knowledge and experience acquired in a specific cultural context?

If US children's difficulty with photo-triggered two-tone reorganization is due to their lack of visual-referential expertise, and such expertise is provided by a culture that trains individuals how to "read" visual symbols such as representational art, maps and models, and writing – then we should find a similar pattern of performance in adults whose culture does not require such expertise. We therefore showed two-tone images and their photo pairs to adult members of the Pirahã, a hunter-gatherer tribe with a sparse visual symbolic culture and limited exposure to modern visual media. Like young children in a modern industrial culture, Pirahã adults have little experience or knowledge of the visual transformation that links a photo and two-tone image. However, they possess both physiologically mature visual systems and a lifetime of experience with complex visual tasks such as hunting and fishing.

Experiment 2A

The procedure and stimuli were modified from those reported in Yoon et al, 2007, to be appropriate for the Pirahã participants.

Participants Adult members of the Pirahã tribe (n = 9, mean estimated age = 30y). The visual acuity of the Pirahã population was tested some years earlier as part of a basic screen for medical services; the population was on the whole normal, with no cataracts and a small incidence of nearsightedness.

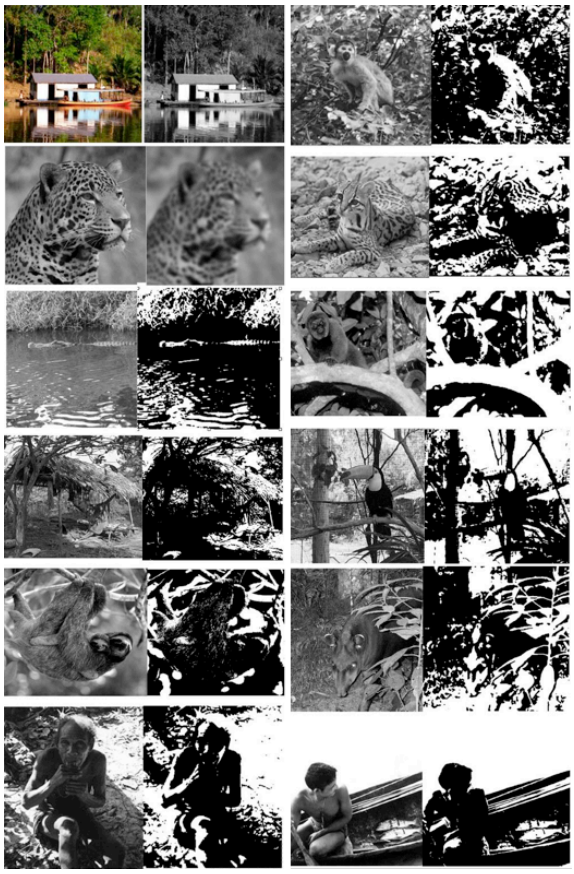

Figure 5. Stimuli used in Experiments 2A, 2B, 2C.

Stimuli Ten two-tone images were created by blurring and thresholding grayscale photographs of animals and individuals found in the Pirahã participants' everyday environment: tribe members, monkeys, a toucan, ocelot, tapir, and sloths. The amount of blur and the black/white threshold point were set independently for each photograph to meet the subjective criteria that the two-tone was hard to recognize without first seeing the photograph from which it was derived ("uncued") but easy to see after seeing the photograph ("cued"). Two other image pairs were tested which did not include two-tones and for which the correspondence was easier to see (blurred jaguar, grayscaled houseboat). These served as warm-up items and to ensure subjects understood the task. Images were printed onto 12x12cm cards (see Figure 5).

Procedure The experimenter sat with the participant at a small table. Each trial consisted of three stages. In stage 1, subjects were shown a two-tone image and asked to indicate their recognition by pointing to the location of the eye or Pirahã person in the picture. Responses were marked by placing a sticker at the indicated locations. Trials in which the target was not initially identified were considered "candidate reorganization trials." These trials were of particular interest as they provided a test of whether an initially unrecognized two-tone image could be successfully reinterpreted after seeing the corresponding photo. These trials proceeded to stages 2 and 3. In stage 2, subjects were shown the corresponding photograph alone and asked to point to the location of the eye or Pirahã person. In stage 3, the two-tone image and photograph were shown side-by-side and the subject was again asked to point to the location of the eyes or person in the two-tone image.

Results and discussion Pirahã participants successfully indicated the target locations (either eye or person) on the non-two-tone practice images without the corresponding photo cue (93%), showing subjects understood the task (Figure 6, blue bars). Uncued two-tone recognition in Pirahã subjects was infrequent (22% of trials). They identified the targets in the corresponding, untransformed photos 87% of the time (Figure 3, left bars). All Pirahã subjects correctly indicated the target on at least 7 of the 10 photos. Data from trials where the Pirahã did not correctly recognize the photo were excluded from subsequent analysis.

Performance was assessed on candidate reorganization trials by calculating the percentage of two-tones recognized after viewing the photo cue out of all two-tones not initially recognized. Pirahã subjects succeeded on candidate reorganization trials only 30% of the time. Two Pirahã subjects never demonstrated perceptual reorganization, and the highest rate of reorganization for any Pirahã individual was 60%.

Experiment 2B

We tested US adults under identical experimental conditions as a control.

Participants Stanford University students, faculty, and staff (n = 8, mean age = 26y).

Stimuli See Experiment 1.

Procedure See Experiment 1.

Results and discussion U.S. control subjects on the same task successfully indicated the target locations (either eye or person) on the non-two-tone practice images without the corresponding photo cue (controls 100%) showing subjects understood the task (Figure 3, middle bars). Controls located the targets successfully in uncued two-tone images on 73% of trials, more frequently than the Pirahã subjects. Controls identified the targets in the corresponding, untransformed photos 100%. Controls consistently showed photo-triggered perceptual reorganization, always (100%) correctly indicating the eye or Pirahã person on previously unrecognized two-tones.

Experiment 2C

Misaligned presentation was used to control for the possibility that the U.S. subjects succeeded on the task not by recognizing the two-tone images, but merely by locating the point on the two-tone in the same location relative to the image frame as the corresponding point in the photograph.

Participants Stanford students (n = 10, mean age = 19y).

Stimuli Images used in Experiments 1 and 2 were cropped by 10% on two adjacent sides (e.g,. top and left) chosen at random, with the constraint that the corresponding two-tone and photo were not cropped on the same two sides. Thus the eye or head was in a different location relative to the image frame in the photo and in the two-tone.

Procedure See Experiment 1.

Results and discussion Control subjects in the misaligned condition, like the controls in the main experiment and in contrast to the Pirahã subjects, showed near perfect performance on candidate reorganization trials (94%), as would be expected if control subjects experienced reorganization, and did not depend solely on a spatial alignment strategy to localize features.

General Discussion

These data suggest at a minimum, that a mature visual system is insufficient to guarantee photo-triggered perceptual reorganization. It is possible that a similarly low rate of photo-triggered reorganization of two-tone images in young US children and Piraha adult are the result of unrelated causes. But the existence of reduced reorganization in an adult population opens up the possibility that developmental failures in perceptual reorganization in young US children may also be explained by a mechanism distinct from visual system maturation. Thus, one possibility is that the results reported here, together with the previous studies on young children (Deloache et al, 1997; Kovacs & Eisenberg, 2004; Yoon et al., 2007), suggest that the very act of bringing our knowledge and experience to bear on perception in the way required for photo-triggered reinterpretation of a two-tone image may be the result of training and experience that is culture-specific – a kind of perceptual literacy.

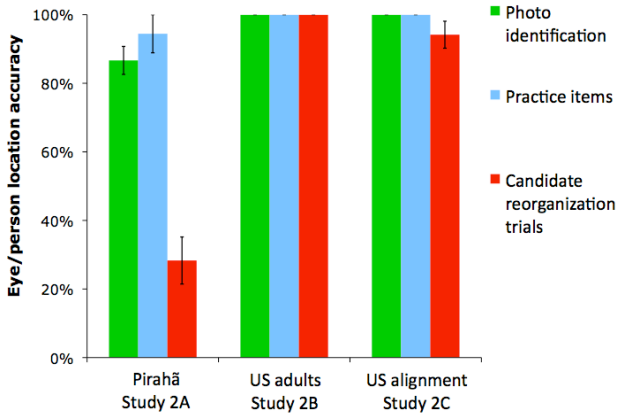

Figure 6: Data summary from Experiments 2A, 2B, 2C.

Culture-specific experience with visual symbols may provide the training required for observers to navigate the dual nature of the two-tone and photo as objects in and of themselves, as well as representations of one another that are mutually informative. Eliminating the need to achieve such dual representation aids inexperienced members of a culture (young US children) in using symbolically corresponding visual representations, such as a scale model, before they acquire expertise in 'reading' visual symbols such as writing and maps (Deloache et al, 1997). Of course, dual representation per se is not impossible for children even much younger than those we tested, depending on the nature of the correspondence. By age two, children infer that a pictorial image corresponds to a real world three-dimensional object even if they have not yet seen such an object. By age three, children are readily able to use a scale model as a "map" to a larger scale room. Therefore, whatever cognitive operation is involved in "pure" dual representation is likely available to children as young as two. Some additional difficulty must explain why it takes an additional year to master dual representation problems involving scale models and a further two or three years when two-tone and photo pairs are involved.

Children's two-tone recognition is somewhat, though not completely, improved when they are given a strong conviction that the two-tone is a physical transformation of a photo image. One idea is that they must understand that

the photo and two-tone should be mapped onto each other in a particular way that preserves spatial relationships even if many local details are mismatched – much as the way a subway diagram preserves the relative relationships between stations while distorting some details about specific distances, and not in the way a Picasso painting can represent a human figure. This may represent a correspondence problem more difficult in degree or kind than that of scale models, where overall spatial layouts are perceptually unambiguous. The two-tone image is ambiguous because the strong edges within the image itself can be grouped in many ways, many of which do not resolve into a meaningful, coherent percept. Unlike unambiguous line drawings and scale models, photo-triggered two-tone reorganization requires a representational override. A conviction that the two highly dissimilar visual objects are transformations of one another is one way to trigger this perceptual override. This practice of perceptual revision, a habit of reinterpreting images when receiving information extrinsic to the image, is part of what may be entrained in a visual symbolic culture and may explain why, by adulthood, a transformation cover story is no longer necessary.

Further research should identify what cognitive skills are entrained by visual symbolic cultures, and how such enculturation may influence other practices of perceptual inference and interpretation necessary for photo-triggered perceptual reorganization.

Acknowledgments

We are grateful to all the participants in these studies: the teachers and children of Bing Nursery School, Stanford University community members, and the Pirahã. Special thanks to Daniel Everett for arranging access to Pirahã, and his prior work characterizing their culture and language. We also thank research assistants who helped with stimulus preparation and/or US data collection and scoring in alphabetical order: Hannah Jaycox, Ben Kessler-Reynolds, Caitlin McLean, Katie O'Neal, Heather Purnell. Special thanks to David Montez and William Beecher for the procurement and modification of the "special machine" used in Experiment 4.

References

DeLoache, J.S., Miller, K.F., & Rosengren, K. S. (1997). The credible shrinking room: Very young children's performance with symbolic and non-symbolic relations. *Psychological Science, 8,* 308-313.

Gregory, R.L. (2001). The Medawar Lecture 2001 knowledge for vision: vision for knowledge, *Philos Trans R Soc Lond B Biol Sci, 360*, 1231-51.

Kovacs, I. & Eisenberg, M. (2004). Human Development of Binocular Rivalry, in *Binocular Rivalry*, D. Alais and R. Blake, Eds. Cambridge, MA: MIT Press, pp. 101-116.

Potter, M.C., Levy, E.I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology, 81*, 10–15.

Yoon, J.M.D., Winawer, J., Witthoft, N., & Markman, E.M. (2007). Striking deficiency in top-down perceptual reorganization of two-tone images in preschool children. *Proceedings of the 6th IEEE International Conference on Development and Learning*.