Hysteresis in the Dynamic Perception of Scenes and Objects

Sonia Poltoratski and Frank Tong
Vanderbilt University

Scenes and objects are effortlessly processed and integrated by the human visual system. Given the distinct neural and behavioral substrates of scene and object processing, it is likely that individuals sometimes preferentially rely on one process or the other when viewing canonical "scene" or "object" stimuli. This would allow the visual system to maximize the specific benefits of these 2 types of processing. It is less obvious which of these modes of perception would be invoked during naturalistic visual transition between a focused view of a single object and an expansive view of an entire scene, particularly at intermediate views that may not be assigned readily to either stimulus category. In the current study, we asked observers to report their online perception of such dynamic image sequences, which zoomed and panned between a canonical view of a single object and an entire scene. We found a large and consistent effect of prior perception, or hysteresis, on the classification of the sequence: observers classified the sequence as an object for several seconds longer if the trial started at the object view and zoomed out, whereas scenes were perceived for longer on trials beginning with a scene view. This hysteresis effect resisted several manipulations of the movie stimulus and of the task performed, but hinged on the perceptual history built by unidirectional progression through the image sequence. Multiple experiments confirmed that this hysteresis effect was not purely decisional and was more prominent for transitions between corresponding objects and scenes than between other high-level stimulus classes. This finding suggests that the competitive mechanisms underlying hysteresis may be especially prominent in the perception of objects and scenes. We propose that hysteresis aids in disambiguating perception during naturalistic visual transitions, which may facilitate a dynamic balance between scene and object processing to enhance processing efficiency.

Keywords: scene perception, hysteresis, object perception, psychophysics, natural images

Visual perception is shaped by the input that one encounters most often. As such, considerable research has focused on the perception of stimuli that are most common in our visual experience: objects, scenes, faces, and bodies. Herein, we consider how the visual system may process two of these stimulus classes, objects and scenes, which are frequently encountered together.

A large body of research has focused on the mechanisms that underlie the processing of scenes and objects; these literatures highlight that these two categories can pose very different problems for the visual system. As described by Kersten, Mamassian, and Yuille (2004), object perception requires the extraction and organization of noisy local image features into objects and surfaces. By some accounts, objects must be separated from the background of the surrounding scene before they can be recognized (Bartels, 2009; Nakayama, He, & Shimojo, 1995; Wolfe, Oliva, Horowitz, Butcher, & Bompas, 2002). In contrast, scene perception has typically been thought to require almost no attention. In fact, several lines of research have considered whether natural scenes are inherently "special" to the visual system because of the speed and ease of their recognition (Cohen, Alvarez, & Nakayama, 2011; Fei-Fei & Perona, 2005; Greene & Oliva, 2009b). Critically, computational models of scene recognition have demonstrated successful scene classification based wholly on low-level image features, without the explicit use of object recognition or segmentation processes (Fei-Fei & Perona, 2005; Oliva & Torralba, 2001; Renninger & Malik, 2004; Torralba & Oliva, 2002).

This ultra-rapid processing may be possible because of the statistical regularities present in scenes. Although object and scene perception both require some degree of invariance to the observer's viewpoint, the larger spatial scale of both indoor and outdoor scenes makes them less prone to large changes during natural viewing (Torralba & Oliva, 2002). Therefore, different types of scene images, such as those of beaches, forests, or cities, have pervasive visual regularities within their category (Oliva & Schyns, 2000), including spatial envelope, constancy, and navigability (Greene & Oliva 2009a; Oliva & Torralba, 2001). Objects, on the other hand, can produce highly variable retinal images depending on their position relative to the viewer. Much research has emphasized the extraction of view-specific or view-invariant features of objects (for reviews, see Biederman & Gerhardstein, 1993; Riesenhuber & Poggio, 2000; Tarr & Bülthoff, 1995; Tarr, Williams, Hayward, & Gauthier, 1998), as well as the extraction of three-dimensional (3D) shape and curvature information (Kourtzi,

---

Erb, Grodd, & Bülthoff, 2003; Todd, 2004). It has been proposed that specific 2D views may be integrated to form a 3D representation (Biederman & Gerhardstein, 1993) or alternatively, that these representations may simply become linked through associative learning (Biederman & Gerhardstein, 1993; Grill-Spector, Kushnir, Edelman, & Avidan, 1999; Kourtzi & Kanwisher, 2001; Tarr & Bülthoff, 1995; for a review, see Bülthoff, Edelman, & Tarr, 1995). These visual representations of objects interact with a rich learned framework for the categorization of objects (for a review, see Palmeri & Gauthier, 2004; Schyns, 1998) in which categories are often defined by features other than pure visual similarity (Goldstone, 1994). Object recognition has been recently modeled as a serial process of "untangling" relevant properties at multiple representational scales (DiCarlo, Zoccolan, & Rust, 2012).

Both scene and object perception require complex visual computation and the integration of visual input with semantic knowledge. Of course, in natural vision, we always see objects in the context of a scene, and most of the scenes we encounter contain at least one easily identifiable object. Although there is strong evidence of distinct processing of objects and scenes, researchers have shown that the recognition of objects is reliably influenced and facilitated by scene context (Biederman, 1972; Palmer, 1975; for a review, see Bar, 2004; Oliva & Torralba, 2007). Similarly, it has also been found that semantically consistent objects improve people's recognition of scenes (Davenport, 2007; Davenport & Potter, 2004). A recent functional MRI study suggested that diagnostic objects contribute to the categorization and neural representation of scenes in the object-selective visual areas, based on the similarity of activity patterns evoked by indoor scenes and specific objects that were diagnostic of those scenes (MacEvoy & Epstien, 2011). This work proposed a mechanism of scene categorization that efficiently combines the representations of several objects, which can facilitate the discrimination of scenes with similar low-level properties. Since objects and scenes almost always co-occur, consideration of how these processes interact is critical for understanding human vision.

Most previous studies have relied on static images to study scene and object perception, as these sufficiently capture many of the complexities of natural vision. However, it is important to consider the dynamic nature of scene and object perception under more natural viewing conditions. For instance, when walking through the woods, one can choose to inspect the richly textured bark of a tree, take a step back and perceive a single redwood, or step back further to take in the entire forest. In fact, most of our waking hours are spent experiencing such continuous and dynamic changes as we actively navigate and explore our immediate environment. Such actions lead to dynamic changes in viewing distance (akin to zooming) and viewing angle as we move our eyes and heads (akin to panning).

Given the computational challenges of object and scene processing, it is likely that we sometimes preferentially rely on one or the other mechanism to maximize its specific advantages. Although certain views of objects and scenes appear unambiguous and may automatically invoke a specific mode of processing, how might observers perceive a gradual transition that changes in spatial scale from a close-up view of a single object to an expansive view that shows the entire scene? The ambiguity in such a stimulus can be demonstrated by the ease with which people can categorize a natural image with verbal labels that correspond to a

specific object or an entire scene. For instance, when considering the image shown in Figure 1A, most people would likely label it as a coffee table, even though the coffee table is pictured in a full-scene context. Similarly, the image in Figure 1B can be readily classified as a living room despite the fact that the scene contains several nameable objects, including the same coffee table. However, if we consider an intermediate image along the transition between Images A and B (i.e., Figure 1C), one might struggle to instantly classify it as either a scene or an object. Hence, the performance of this classification task can reveal an ambiguity in object and scene perception that may be present in our daily lives.

Rather than alternating between two perceptual frameworks in such periods of ambiguity, it may be adaptive for the visual system to consider spatiotemporal continuity and recent visual memory to preserve the seamless nature of perception. In this way, ambiguity during the dynamic perception of objects and scenes may be resolved by the mechanism of hysteresis. In studies of vision, hysteresis is defined as a mechanism by which prior perception influences subsequent judgment. Many natural and artificial systems display hysteresis; it was originally identified in ferromagnetics and has been reported in fields including thermodynamics, biology, and economics (Angeli, Ferrell, & Sonntag, 2004; Bertotti, 1998; Blanchard & Summers, 1986; Jiles & Atherton, 1986). Artificial systems are often designed to have hysteresis to minimize unwanted rapid switching around a threshold; these include the Schmitt trigger in electronics and many modern thermostats
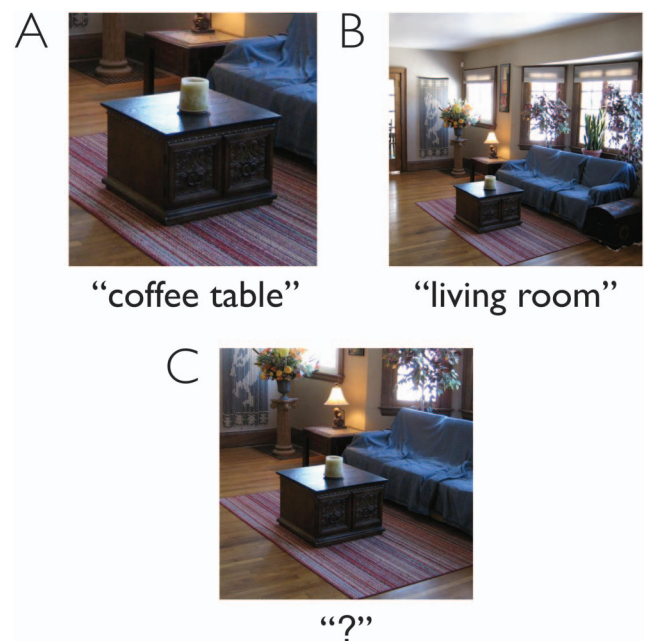

Figure 1. A labeling task that hints at a preferential reliance on either object or scene processing in some cases, and ambiguity in "object-hood" or "scene-hood" in others. The image in Panel A is easily classified as a coffee table, although it is presented in full context. The image in Panel B is likewise easily classified as a living room, although it contains several objects. The image in Panel C, which lies along the perceptual transition between A and B, is more difficult to classify as either a coffee table or a living room. See online article for color version of this figure. © Greg Habermann. Published under a Creative Commons Attribution license.

(Brokate & Friedman, 1989; Schmitt, 1938). Similarly, perceptual decisions about an ambiguous, changing stimulus often exhibit the influence of recent visual experience. Hysteresis has been reported in studies of stereopsis (Buckthought, Kim, & Wilson, 2008; Fender & Julesz, 1967; Julesz, 1974), binocular rivalry (Buckthought et al., 2008), motion perception (Hock, Bukowski, Nichols, Huisman, & Rivera, 2005; Hock, Kelso, & Schöner, 1993; Williams & Sekuler, 1986), the perception of letters emerging from visual noise (Kleinschmidt, Buchel, Hutton, Friston, & Frackowiak, 2002), the identification of line-drawn objects (You, Meng, Huan, & Wang, 2011), and the recognition of facial emotional expressions (Sacharin, Sander, & Scherer, 2012). In a seminal study, Williams and Sekuler (1986) studied hysteresis in the perception of random-dot motion. The coherence of motion steadily ramped up and then down, causing observers to alternate between seeing completely random motion and upward flow. The authors found that the reported boundary between these two percepts was dependent on the observer's recent perceptual history: greater coherence was needed for observers to report upward flow in the initial "ramp-up" portion of the trial than for them to report random motion in the subsequent "ramp-down" phase. The authors identified this as hysteresis in the perception of motion direction and modeled this behavior by assuming cooperative neural interactions between units with similar directional tuning and inhibitory interactions between neurons with different tuning. Subsequent functional neuroimaging work by Kleinschmidt et al. (2002), in which they measured neural substrates of hysteresis in the detection of letters with changing contrast, has implicated not only regions in the visual cortices but also frontal and inferior parietal regions. This finding highlights the interaction of vision, attention, and decision making in hysteresis, even when responses to simple features are measured.

We hypothesized that naturalistic viewing of objects and scenes would be specifically prone to hysteresis. This hypothesis is made on the basis of several aspects of our visual experience of objects and scenes: (a) their processing is separable (neurally and as a function of spatial scale), (b) their processing is cognitively demanding (even in the case of simple gist extraction), and (c) they almost always co-occur within a single environment.

Although objects and scenes certainty interact on the level of semantic context (Biederman, 1972; Palmer, 1975; for a review, see Bar, 2004; Davenport, 2007; Davenport & Potter, 2004; Dilks, Julian, Paunov, & Kanwisher, 2013; Oliva & Torralba, 2007), these two stimuli are thought to involve neurally separable processes (Epstein, 2008; Epstein & Kanwisher, 1998; Grill-Spector et al., 1999; Park & Chun, 2009). When both scene and object information is present in an image, research has suggested that processing is made more efficient by a sequential "global-to-local" sweep, in which global properties are encoded before local ones (Cesarei & Loftus, 2011; Navon, 1977; Schyns, 1998; Schyns & Oliva, 1994; Shulman & Wilson, 1987). This can also be described as "coarse-to-fine" processing, as spatial frequency (SF) has been found to sufficiently characterize the global-to-local distinction in natural images (Schyns, 1998; Shulman & Wilson, 1987). Scene processing has often been described in terms of global or low SF information (Greene & Oliva 2009a; Oliva & Torralba, 2001; Renninger & Malik, 2004; Torralba & Oliva, 2002; but see Vogel, Schwaninger, Wallraven, & Bülthoff, 2006), while objects are more informative at a higher SF, requiring local processing (Ker-

sten et al., 2004; Parker, Lishman, & Hughes, 1996). Of course, this frequency-based selection must flexibly accommodate task demands (Collin & Mcmullen, 2005; Schyns, 1998) and thus might not always take precedence over the encoding of local objects. Bar (2004) has proposed a cortical feedback model of this visual analysis process, in which initially processed global information influences subsequent fine-scale processing (see also: Evans & Treisman, 2005).

It should be noted that both global and local visual processing require the use of limited attentional resources. In fact, even ultra-rapid scene recognition, which was traditionally thought to require essentially no attention (Fei-Fei, VanRullen, Koch, & Perona, 2002), shows evidence of processing costs when more than one scene is presented simultaneously (Rousselet, Thorpe, & Fabre-Thorpe, 2004); scene perception can also be hindered by having to perform a secondary, demanding nonscene task (Cohen et al., 2011). One can imagine that in cases excluding facilitation by semantic context—which, at 300–400 ms, occurs on a similar time course as purely verbal consistency effects (Ganis & Kutas, 2003)—simultaneous object and scene judgments should incur a typical dual-task cost (Pashler, 1994).

However, most of the scenes we encounter contain at least one object, and objects usually occur in a visible scene context. When both scenes and objects are in view, the visual system may prioritize one type of processing over the other to allocate limited cognitive resources to demanding operations. To do this efficiently, the system should minimize rapid switching between object and scene perception at intermediate views. This can be accomplished through hysteresis, such that one's sensory memory influences whether scene or object processing is engaged at any given moment. Although hysteresis is likely to contribute to many aspects of perception, in this study we addressed the specific hypothesis that this mechanism might play a particularly prominent role in the perception of natural scenes and objects. This hypothesis is largely motivated by our natural experience with objects and scenes, which likely requires switching between object and scene processing yet appears perceptually coherent.

In the current study, we investigated the interaction of object and scene processing under dynamic naturalistic conditions to evaluate the effects of prior viewing history on perceptual judgments. We did this by constructing simple movies that consisted a sequence of images that smoothly transitioned between a canonical object view and a canonical scene view (see Figure 2) and asking observers to report their subjective categorical boundary between "object" and "scene." We hypothesized that perceptual classification of the transition from object-to-scene or scene-to-object would be subject to hysteresis. That is, an image sequence would be labeled as an object for a longer period of time when the participant started viewing at the object end point, whereas the converse would be true when the participant started with viewing of the scene end point. If the scene/object classification was based solely on properties of the image and not on perceptual history, then we would expect judgments between the object-first and scene-first movies to only differ in the range of human reaction times (by a few hundred milliseconds in each direction). By using this dynamic object/scene categorization task, we sought to quantify the effect of recent viewing experience on the classification of objects and scenes. Moreover, we sought to quantify the effect of hyster-
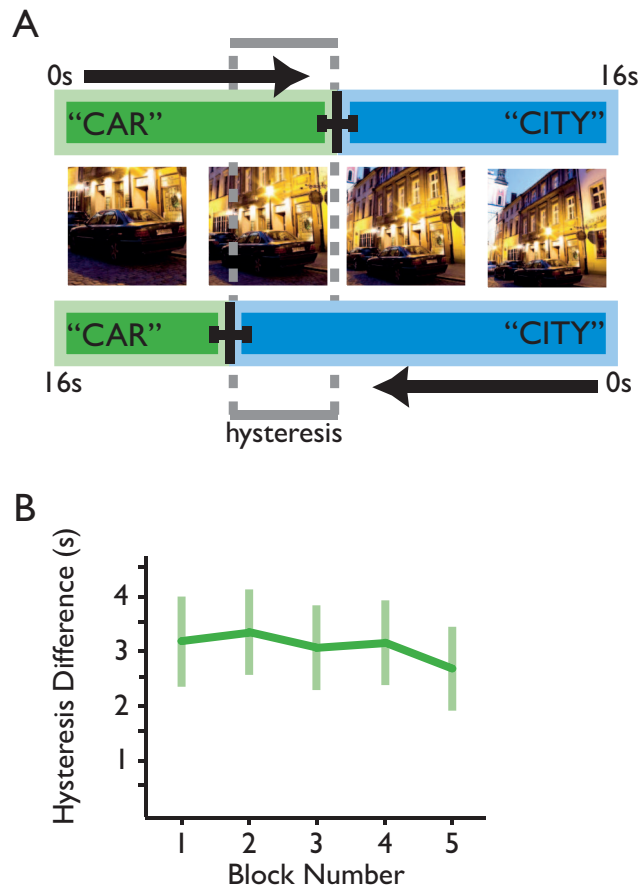
A

B

Figure 2. Results from Experiment 1: Smooth "movie-style" presentation of sequences between objects and scenes. Panel A: Data from all subjects across all movies are overlaid on a graphic depiction of a single trial for clarity. Time is plotted horizontally, and each frame of the chair-to-beach sequence represents 4 s. Accordingly, Frames 30 (near the object end point, 2 s into the movie), 90, 150, and 220 (near the scene end point, 14 s into the movie) are shown. The categorical boundary judgment in the object-to-scene direction is plotted above the graphic, and the judgment in the scene-to-object below the graphic. The hysteresis difference between the two directions is 3.07 s ($p < .005$). Panel B: The size of the hysteresis effect is plotted for each of the experimental blocks. No effect of block was found ($p = 0.23$), demonstrating that familiarity with the sequences did not mediate the effect. See online article for color version of this figure. Images courtesy of Kuba Cichocki Photography.

esis and to determine its specificity to both the stimulus and the task.

In Experiment 1, we found that hysteresis indeed occurs when observers classify objects and scenes in dynamic transition. In Experiments 2 and 3, we asked whether the effect hinged on low-level motion cues, unidirectional progression through the image sequence, or passive viewing. In Experiment 4, we asked whether uncertainty in the decision could cause observers to respond in a way consistent with hysteresis and tested for the presence of hysteresis after taking into account the potential effects of uncertainty and of categorization. In Experiment 5, we provided observers with a financial incentive to avoid hysteresis in their categorization judgments. Even with such incentives in place,

observers exhibited clear effects of hysteresis before they adopted alternative cognitive strategies to minimize the impact of prior viewing history. Finally, in Experiment 6, we evaluated a new experimental paradigm that involved smoothly blending between two discrete images by dynamically adjusting their relative contrasts; although less naturalistic, these image blends allowed us to compare the magnitude of hysteresis between distinct categories of stimuli.

In all of these experiments, we find a strong and persistent effect of hysteresis that cannot be explained by the specific parameters of our stimuli or by uncertainty in the decision. Observers were unable to eliminate the hysteresis effect even when highly motivated to do so. In the contrast ramping experiment, hysteresis was also markedly greater for transitions between congruent objects and scenes than between other stimuli pairs. We suggest that hysteresis, a mechanism for reducing ambiguity at perceptual boundaries, has a particularly profound influence in the perception of objects and scenes.

Experiment 1: Hysteresis in Object/Scene Categorization

In the first experiment, we measured scene/object classification of dynamic movies that portrayed a gradual transition from a zoomed-out view of a real-world scene to a zoomed-in view of a view of central object within that scene. We constructed 16 movies, and each movie lasted 16 s. Movies were presented in both possible directions, displaying transitions from object to scene and from scene to object. Observers were instructed to make a single key press to report their subjective boundary between which portion of the movie primarily depicted an object and which primarily depicted a scene. Since each movie stimulus was shown in both object-to-scene and scene-to-object directions, the responses between these two directions could be compared to test for effects of perceptual hysteresis.

This experiment allowed us to measure the effect of viewing history on the classification of a dynamic image sequence that gradually transitioned between a canonical scene and a canonical object image. We predicted that observers would perceive the intermediate portion of the movie as somewhat ambiguous and that the observers' recent perceptual history would influence the classification judgment. Specifically, we hypothesized that each image sequence would be classified as an object for a considerably longer duration when it was viewed in the object-to-scene direction than in the scene-to-object direction, and vice versa.

Method

Participants. Ten undergraduate students (two males and eight females) from Vanderbilt University (ages 18–21 years) participated in this experiment for course credit. All had normal or corrected-to-normal visual acuity and reported having normal color vision. The Vanderbilt University Institutional Review Board approved the study.

Stimuli and apparatus. For these experiments, we constructed sixteen 240-frame object-to-scene sequences, which were made by repeatedly sampling a square-sized region from a single high-resolution image at multiple spatial scales and rescaling the sampled region to a 500 × 500 pixel image. These

movie-style stimuli were presented at a rate of 15 frames per second. We initially selected images that depicted a scene with one or more central objects. Images were then cropped to contain the most zoomed-out view of the scene image, which served as one end point of the image sequence. From this, we defined the most zoomed-in object view, which consisted of a region of interest that was approximately 25% of the size of the scene image. The object region, thus defined, served as the other end point of the transition. The intermediate frames were generated by a Matlab (MathWorks, Natick, MA) script that cropped inward from the scene image to the object image in 240 iterations, using linearly spaced steps to adjust the size of the sampling window. Finally, each of these frames was resized to 500×500 pixels to ensure a constant image size throughout the movie. Stimuli were presented at the center of a computer monitor on a uniform black background. At the estimated viewing distance of 55 cm, the display subtended approximately 17.6° of visual angle.

Visual presentation and response recording were carried out using Matlab and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) on an Apple Mac Mini computer (Apple, Inc., Cupertino, CA) with a 2-GHz Intel Core Duo processor (Intel Co., Santa Clara, CA) and a Sony Trinitron CRT monitor (1152×870 resolution at 75 Hz; Sony Corp., Tokyo, Japan).

Procedure. Observers were seated in a private, quiet testing room and were given a verbal explanation of the task by the experimenter, as well as a visual schematic of the trial and experiment structure. On each trial, participants were provided with the appropriate object and scene labels and asked to press a key when they perceived a transition between the two (for example, a trial could read "Press a key when the movie changes from COFFEE TABLE to LIVING ROOM").

Each subject performed five blocks of trials (32 trials per block), viewing each transition in both the object-to-scene and scene-to-object direction in each block. The first direction that a particular observer saw each transition was counterbalanced across blocks, so that eight movies were shown in the object-to-scene direction in the first half of each block, and the other eight movies were shown in the scene-to-object direction first. Moreover, the first direction of these blocks of eight movies was counterbalanced across participants. Movies were presented in a randomly shuffled order within each block, subject to the previously described constraints; also, the same movie stimulus could not appear twice in a row. The entire experiment consisted of 160 trials and required approximately 50 min to complete. We collected the first key press on each trial that the observer made to mark her or his boundary between object and scene and recorded how long into the movie this occurred.

Results and Discussion

We found a large and consistent effect of perceptual hysteresis when participants viewed a smooth visual transition from scenes to objects and also objects to scenes. The average time of response across subjects is plotted in Figure 2, with images of a representative movie shown alongside to depict the stimulus that was shown at 2, 6, 10, and 14 s. A repeated-measures analysis of variance (ANOVA) revealed a main effect of transition direction, $F(1, 9) = 14.96$, $p < .005$, partial $\eta^2 = .624$. This effect was on average 3.07 s ($SE = 0.800$ s) and in the direction predicted by

hysteresis. That is, the movie stimulus was judged to depict an object for a longer period of time when the movie started at the object image and zoomed out toward the scene than when the same movie sequence was shown in the opposite direction, and the response time difference of 3.07 s was much larger than one would expect from simple reaction time delays. There was no significant effect of block within the experiment, $F(4, 36) = 1.475$, $p = .230$, partial $\eta^2 = .141$, nor interaction of block with direction, $F(4, 36) = .902$, $p = .473$, partial $\eta^2 = .091$, indicating that this effect was consistent across repeated presentations of the stimuli and not unique to the first presentation of a novel, dynamic stimulus.

Each movie portrayed very different objects and scenes, and the stimulus set was quite variable in terms of degree of visual clutter, percentage of area occupied by the selected objects, and many other features. Several representative examples of the stimuli are shown in this article's figures; for the full set of object and scene endpoints of the 16 movies shown in Experiments 1–5, please contact the corresponding author (sonia.poltoratski@vanderbilt .edu). Our key behavioral measure should not be sensitive to such stimulus variability, as the exact same movie was shown in both object-to-scene and scene-to-object directions. However, it is still possible that a small subset of the movie stimuli drove the effect. To explore this, we plotted both the frequency distribution of the estimated hysteresis effect for all movies and participants (Figure 3A; data from 10 participants for each of 16 movies, a total of 160 observations) and an estimated density function to visualize the hysteresis effect for each movie (Figure 3B). The analyses indicated that the hysteresis effect was present in all movies, with no apparent outliers.

We thus found a strong and consistent effect of perceptual hysteresis when observers reported their subjective boundary between object and scene during viewing of these dynamic movie stimuli. This 3.07-s effect exceeded the predictions of simple reaction time differences and was found to be consistent across a highly variable stimulus set. Our experiment compared identical image sequences that varied only in the directional order in which they were shown; thus, we can suggest that hysteresis is likely driven by the observer's recent viewing history and not by a physical property of the movie stimuli. However, the movie stimuli used in Experiment 1 did produce strong motion cues. We conducted Experiment 2 to determine whether these motion cues were necessary for the hysteresis effect.

Experiment 2: Removal of Dynamic Motion Cues

In Experiment 2, we minimized dynamic motion cues by presenting the stimuli as a sequence of static images with intervening blank periods, rather than as a smooth "movie" depicted by a continuous series of images. The intervening blank periods lasted for 300 ms, which is long enough to disrupt basic motion perception arising from short-range apparent motion (Burr et al., 1984; Burr & Ross, 1986; Nakayama & Silverman, 1984). We predicted that the effects of hysteresis reflected a tendency for object- and scene-specific processing to persist over time, and as such, the disruption of apparent motion should not impair this hysteresis effect.
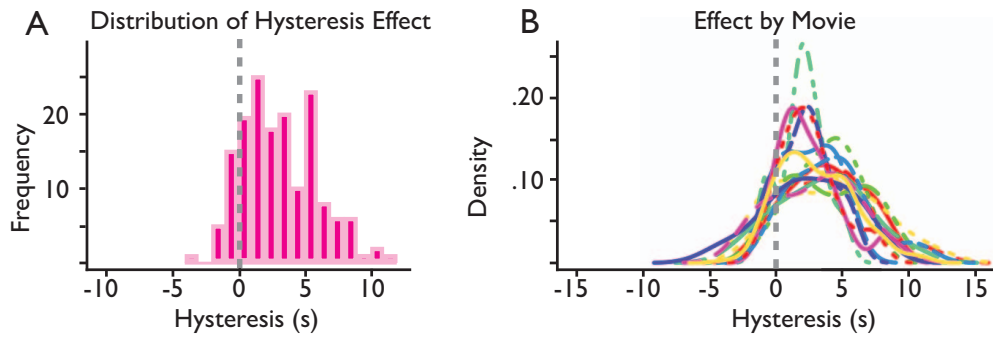
Figure 3. Distributions of hysteresis effects in Experiment 1. Panel A: Histogram of the length of hysteresis for each movie and each observer (160 total data points). Positive values indicate a temporal difference consistent with hysteresis. Panel B: Density plots of the hysteresis effect for each of the 16 movies overlaid on a single axis. Each curve represents the hysteresis effect for a single movie across the 10 observers. It is clear that the average hysteresis effect is not driven by any apparent outliers and is present in all sequences used. See online article for color version of this figure.

Method

Ten new observers (four males and six females, ages 18–20 years) from the same subject pool used in the first experiment participated in Experiment 2. The design was identical to Experiment 1 except that instead of showing all 240 frames of image sequences used in Experiment 1, we selected 33 frames from evenly spaced intervals for each image sequence. We changed the rate of presentation to 2 frames/second (from 15 frames/s in Experiment 1) while preserving the 16-s trial length used previously. Each image was presented for 200 ms each, followed by 300 ms of a blank screen. The duration of this blank interval was sufficiently long to disrupt the perception of low-level apparent motion (Burr & Ross, 1986; Burr, Ross, & Morrone, 1986; Nakayama & Silverman, 1984). Stimuli were presented at the center of the screen on a uniform gray background.

Results and Discussion

When scene/object transitions were viewed as sequences of static images with minimal low-level motion cues, we observed prominent hysteresis effects that were in fact larger than those found with smooth, movie-style transitions. The results of Experiment 2 are shown in Figure 4, depicted with frames sampled from another image sequence.

A repeated-measures ANOVA revealed a main effect of movie direction on time of the reported object/scene transitions, $F(1, 9) = 50.995$, $p < .001$, partial $\eta^2 = .850$, with no effect of block number, $F(4, 36) = .809$, $p = .528$, partial $\eta^2 = .082$. The duration of the hysteresis effect across subjects was 5.54 s ($SE = 0.774$), which was highly statistically significant, $t(9) = 7.147$, $p < .001$, Cohen's $d = 2.25$. Block number interacted with sequence direction, $F(4, 36) = 2.84$, $p < .04$, partial $\eta^2 = .240$, but there is no significant linear trend in this interaction, $F(1, 9) = .556$, $p = .475$, partial $\eta^2 = .058$; the individual hysteresis differences in each block were: 5.92 s, 5.64 s, 5.04 s, 5.04 s, and 5.89 s. We performed an independent samples t test to contrast the hysteresis effects in Experiments 1 and 2, which revealed that the hysteresis effect in Experiment 2 was significantly longer, $t(18) = 2.174$, $p < .05$, Cohen's $d = 0.982$.

Thus, we found that dynamic motion is not necessary to obtain the hysteresis effect; in fact, disrupting low-level motion cues in the stimuli seemed to increase the behavioral difference between the object-to-scene and scene-to-object directions. There are several possible explanations for this finding. One might note that the interrupted presentation of successive images resembles the "flicker" task often used to demonstrate the phenomenon of change blindness (Rensink, O'Regan, & Clark, 1997). According to this account, the use of the flickering display may impede the participants' ability to detect changes from one image to the next, and as a consequence they may take longer to notice that the movie stimulus has crossed their subjective boundary between object and scene. Alternatively, the increase in hysteresis effect might be
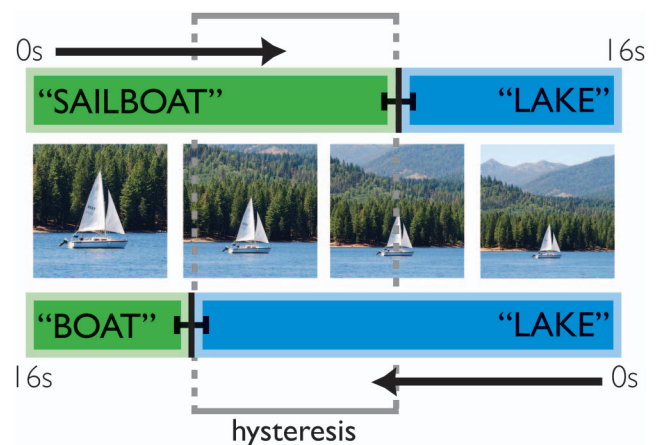

Figure 4. Results of Experiment 2: Removal of dynamic motion cues. Again, results across all observers and transitions are overlaid on a depiction of a single trial for clarity, and the images shown at 2 s, 6 s, 12 s, and 14 s illustrate the degree of change across 4-s intervals. The hysteresis effect between the object-to-scene (top bar) and scene-to-object (bottom bar) presentation directions was 5.54 s ($p < .005$). This effect was significantly larger than the 3.07-s difference found in Experiment 1 ($p < .05$). See online article for color version of this figure. Images courtesy of Kanoa Helms/Daily Venture.

explained by the relative paucity of visual information in Experiment 2 compared with Experiment 1. In Experiment 2, we presented only 33 evenly spaced frames drawn from the full, 240-frame sequence used in Experiment 1. In total, these 33 frames comprised 6.6 s of visual stimulus, compared with 16 s in Experiment 2. If we assume that a certain amount of visual information is needed to make the decision of scene versus object, it may be that this visual information takes relatively longer to accumulate in the case of Experiment 2.

Finally, research has demonstrated that when multistable stimuli are shown using intermittent presentation, the observer's perception tends to stabilize on a certain perceptual interpretation (Leopold, Wilke, Maier, & Logothetis, 2002; Maier, Wilke, Logothetis, & Leopold, 2003; Pearson & Brascamp, 2008). Multistable stimuli, such as binocular rivalry displays, the Necker cube, or a rotating structure-from-motion sphere, spontaneously vary in their appearance over time. However, presenting these stimuli intermittently sharply reduces the frequency of perceptual reversals, stabilizing a single percept for as long as several minutes. These effects of perceptual stabilization are believed to reflect a form of sensory memory.

If the intermittent presentation of our image sequences introduces a similar perceptual ambiguity, then the increased hysteresis effect with intermittent presentation may be related to the stabilization phenomenon discussed by Maier et al. (2003) and Pearson and Brascamp (2008). The latter work highlights that the persistence of a given percept reflects a memory trace of certain low-level image properties that builds over exposure time, rather than a simple snapshot of the last-seen image. It may be that the increased hysteresis seen in Experiment 2 reflects the contribution of similar features and perceptual history.

Regardless, the results of Experiment 2 clearly indicate that motion cues are not necessary for eliciting a hysteresis effect for object-to-scene transitions. In fact, the duration of the hysteresis difference in category judgment is increased when the movie stimulus is interrupted. Further work is needed to determine the cause of this increase, as it may reveal mechanisms or features that enhance the hysteresis effect.

Experiment 3: Self-Paced Progression

The goals of Experiment 3 were to determine whether the hysteresis effect would persist if observers had active control over their progression through the sequence of images and whether hysteresis depends on unidirectional progression through this sequence. The pronounced effects of hysteresis observed in Experiments 1 and 2 occurred under conditions of passive viewing. Although the duration of these effects greatly exceeded the magnitude predicted by simple response time, a potential concern is that brief lapses of attention during passive viewing of these image sequences might have magnified our estimates of the hysteresis effect. To address these potential concerns, we had observers perform an active version of this task in which they advanced through the image sequence of their own volition.

We allowed observers to self-advance through each sequence of images by pressing different keys to zoom in toward the object view or zoom out toward the scene view. A slider appeared underneath the transition stimulus and reflected the key-presses. In Experiment 3A, observers could advance unidirectionally from

one end point of the image sequence to the other, at the pace of their choosing. This design was largely equivalent to that of Experiment 1 but provided observers with active control over the progression through the image sequence. In Experiment 3B, observers could proceed through the image sequence in both forward and reverse directions.

We hypothesized that the hysteresis effect is caused by the persistence of an observer's recent experience of either object or scene perception. This experience would remain uninterrupted when observers could only move through the transition in one direction (Experiment 3A) but would change when bidirectional progression was possible (Experiment 3B). Accordingly, we expected that the hysteresis difference in categorizing the object/scene transition would be eliminated in Experiment 3B but persist in Experiment 3A.

Method

We recruited 24 new observers from the Vanderbilt University community to participate in these experiments for monetary compensation. Twelve observers (four males and eight females, ages 19–21 years) took part in Experiment 3A, and 12 (four males and eight females, ages 19–27 years) took part in Experiment 3B.

The same 16-movie stimuli (240 frames/movie) were used for this experiment and were presented centrally on a uniform black screen. Underneath the central stimulus appeared a graphic slider: a white line representing the entire image sequence, and a cyan square representing the current frame position. On each trial, the object and scene labels for the transition appeared on the left and right side of the slider, respectively. Slider Position 1 always corresponded to the object end point and Position 240 to the scene end point.

In Experiment 3A, each trial started at either Frame 1 or Frame 240. This was done to hold constant the length of each image sequence and to minimize potential confusion about the allowed direction of motion. In Experiment 3B, each trial started at one of four possible slider positions: Frame 1, Frame 60, Frame 180, and the last Frame 240. Frame 120, the absolute midpoint of the image sequence, was not used to avoid the impression that an objective categorical boundary existed at the middle of the sequence. Observers were presented with each image sequence once in each of four blocks; the order of the starting points for these image sequences was counterbalanced across observers. There were 64 trials in total, and their order within each block was randomized.

Observers could move through the transition using the keyboard: keys F and J advanced at a rate of one frame per key press, and keys D and K at a rate of five frames per key press. After receiving a verbal explanation by the experimenter and a visual schematic of the task, observers were instructed to identify the categorical boundary between which portion of the image sequence they thought depicted an object and which portion depicted a scene. In Experiment 3A, observers could only advance the image sequence away from the starting position. They were instructed to press the space bar to record their boundary but to then continue through the sequence until they reached the other end point. This was done to ensure that after the first block, no image in the movie sequence would remain novel, so that familiarity would not be a confounding factor in this version of the experiment. In Experiment 3B, observers could advance through the

image series in either direction and were instructed to press the space bar to record their categorical boundary and end each trial. These experiments were completely self-paced; all observers spent between 20–50 min on the task.

Results and Discussion

The results of Experiment 3A are shown in Figure 5A. When observers could advance through the transition in a single direction but at their own pace, the effect of hysteresis remained highly reliable in the object/scene classification task. This is evident from the difference between the boundaries reported when observers started at the scene end point compared with the object end point, which was on average 28.65 frames. This difference is in the direction predicted by hysteresis, reflecting a persistence of the starting percept. In the movie-style presentation used in Experiment 1, a difference of 28.65 frames would correspond to 1.91 s of hysteresis; however, since observers moved through the transitions at variable rates, it is difficult to compare the length of the hysteresis effect between these experiments. The average trial duration in Experiment 3A was 12.40 s (SE = 1.64 s). A repeated-measures ANOVA revealed a significant effect of starting position, $F(1, 11) = 7.560$, $p < .020$, partial $\eta^2 = .407$, but no effect of block number, $F(3, 33) = .314$, $p = .815$, partial $\eta^2 = .028$, nor any evidence of an interaction between the two factors, $F(3, 33) = .350$, $p = .789$, partial $\eta^2 = .031$. These findings indicate that the hysteresis effect remained consistent throughout the experiment and did not change reliably across blocks.

However, when observers had the opportunity to advance through the image sequences bidirectionally in Experiment 3B, a

very different pattern of results emerged. Observers favored moving bidirectionally through the sequence: on 94.1% of all trials (SD = 7.15%), they advanced past their eventual boundary location before reversing direction to report their decision. With the opportunity for bidirectional progression, the hysteresis effect disappeared: there was no effect of starting position on observers' judgment of the categorical boundary between object and scene (see Figure 5B).

A repeated-measures ANOVA showed no effect of starting position, $F(3, 33) = 1.59$, $p = .210$, partial $\eta^2 = .126$, no effect of block, $F(3, 33) = .132$, $p = .940$, partial $\eta^2 = .012$, and no evidence of an interaction effect, $F(9, 99) = .753$, $p = .660$, partial $\eta^2 = .064$, for the boundary judgment. For each individual subject, we also performed a paired t test to compare the average boundary reported for Starting Positions 1 and 240, the object and scene end points, which corresponded to the starting positions used in Experiments 1, 2, and 3A. This difference was significant for only one of the 10 observers at $p = .049$ (uncorrected for multiple comparisons), again suggesting no reliable difference in reported boundary across starting positions. On average, observers viewed each trial for 19.01 s (SE = 3.27 s).

From these results, we see that self-paced advancement alone cannot eliminate hysteresis (Experiment 3A) but that bidirectional movement through the image sequence can (Experiment 3B). When subjects could advance through the transition at a speed and direction of their choice, they converged on a boundary between scene and object that was statistically comparable to the midpoint of the two direction-specific responses recorded in Experiments 1 and 3A, $F(2, 18) = 2.73$, $p = .092$, partial $\eta^2 = .144$.

These two experiments suggest that hysteresis in dynamic object/scene classification is not a consequence of attention lapses during passive viewing of the sequences. Experiment 3A demonstrates that response time limitations cannot account for the effect, as observers in that experiment had unlimited time to view each frame and make a response. We can also infer that in Experiment 3B, bidirectional progression affected the consistency of the observer's perceptual history prior to making the categorization and that some degree of this uninterrupted perceptual history is necessary for hysteresis.

Experiment 4: Uncertainty in the Decision

In the preceding experiments, we have demonstrated an enduring effect of recent perceptual history on the classification of gradual transitions from objects to scenes. This is manifested in a difference in the reported object/scene boundary when observers view the same image sequence from scene-to-object and from object-to-scene. However, in doing so, we have assumed that this task involves only two possible perceptual states: that is, each time point of the sequence can be classified as either a picture of an object or as a picture of a scene.

An alternative explanation of this difference in reported boundaries is the presence of uncertainty in the classification task, as is illustrated in Figure 6. This figure depicts a simple model of uncertainty and its potential effects on the response time of category judgments. It is possible that a period of uncertainty occurs during the middle portion of each image sequence: perhaps the observer has decided that the movie no longer depicts a flower (the object) but has not yet amassed
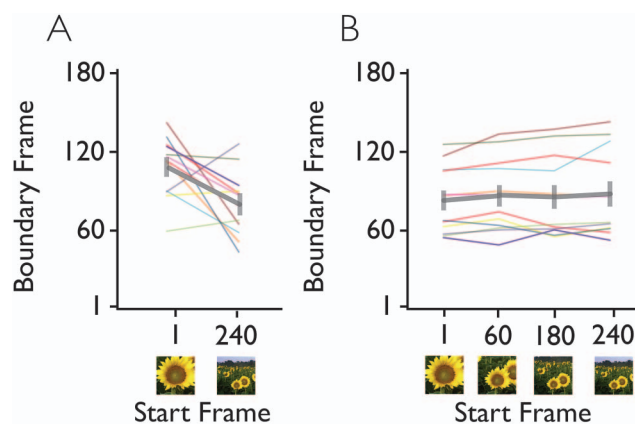

Figure 5. Panel A: Results of Experiment 3A are shown: scene/object boundary using unidirectional active progression. Boundary judgments are plotted for each subject, and averages are shown in gray. There was a significant difference between reported boundaries when observers started the trial at the "object" end point (Frame 1) and the "scene" end point (Frame 240). This difference was in the direction predicted by the hysteresis hypothesis, confirming that lapses in attention during passive viewing cannot explain the effect. Panel B: Results of Experiment 3B are shown: scene/object boundary using a bidirectional active progression. Allowing observers to reverse their motion through the sequence eliminated hysteresis from the boundary reports. There was no significant effect of starting position across subjects ($p = .21$). See online article for color version of this figure. Images courtesy of Stephen L. Tabone Nature Photography (STabone.com/StephenLTaboneBlog.com).
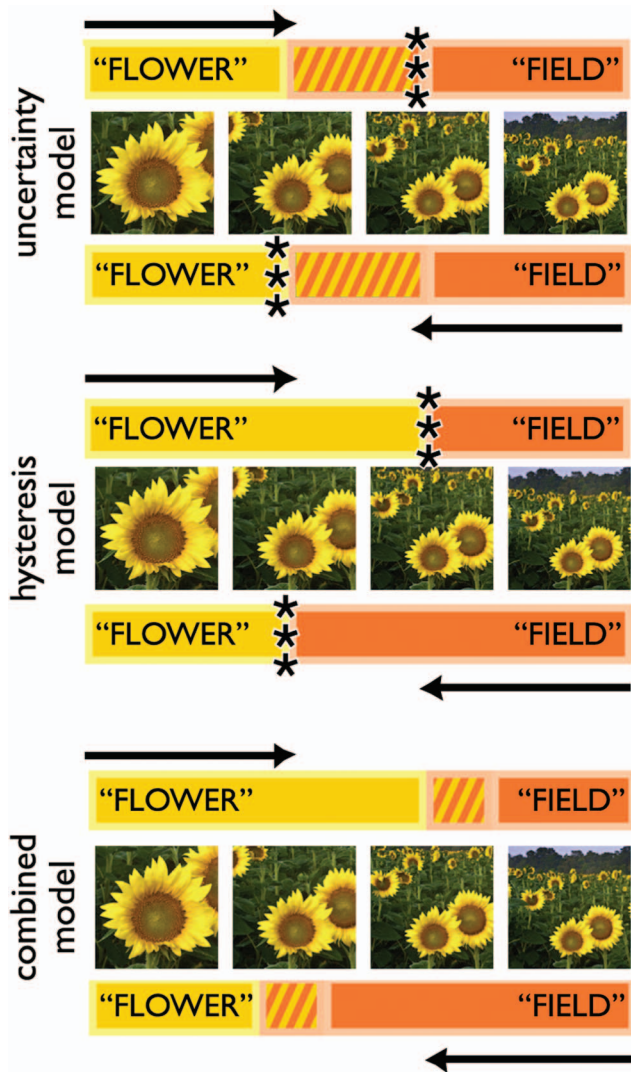
Figure 6. Models of uncertainty and hysteresis are shown. In these models, periods in which the sequence is classified as a flower are shown in the lighter portion on the left and periods in which the sequence is classified as a field are shown in the darker portion on the right. In the uncertainty model, there is an intermediate period, shown in stripes, during which the observer does not classify the sequence as either a flower or a field. When an uncertain observer is asked to report "When the movie changes from flower to field," he or she may wait to accumulate evidence and make reports after the period of uncertainty, which would coincide with the reports predicted by hysteresis. In the combined model, a period of uncertainty is present, as is a hysteresis period, during which a portion of the sequence is classified with certainty as either a flower or a field depending on the viewing direction. See online article for color version of this figure. Images courtesy of Stephen L. Tabone Nature Photography (STabone.com/StephenLTaboneBlog.com).

enough evidence to classify the movie as a field (the scene) and will not make a key press until after this threshold has been reached. This is important for the current study because, as shown in Figure 6, the response pattern predicted by the uncertainty model may look identical to that predicted by hysteresis. Specifically, if the periods of object and scene perception are

constant between the two movie directions, but there is a "gap" of uncertainty between them, an observer may not respond that the transition has changed until he or she is certain. This would result in the same apparent lag in reporting the boundary that we have interpreted as hysteresis. Of course, it may be that both uncertainty and hysteresis are present in the decision, as illustrated in the "combined model" of Figure 6. In this case, there is both a period of uncertainty, for which the sequence is classified as neither an object nor a scene, and a period of hysteresis, for which judgments of object and scene are shifted, depending on the presentation direction of the sequence.

In Experiment 4A, we used an alternative reporting method to quantify the period of uncertainty in the object/scene classification task. We changed the wording of the task so that the classification would not be assumed to be a two-alternative choice between object and scene. We presented the same image sequences, but only required observers to classify them as "scene" or "not scene" and as "object" or "not object." We did this by asking observers to press a key when the movie "stops being a ____" in one block, and, in another block, when the same movie "starts being a ____." We refer to this classification task as the *single-category* decision, as each trial prompted observers to consider either "object-hood" or "scene-hood" in the movie but not both concurrently.

This experimental design allowed us to separately estimate the duration of both the uncertainty effect and hysteresis effect. Uncertainty was estimated by comparing responses between the two single-category blocks for a single presentation direction of a sequence. If uncertainty was present in the object-to-scene presentation direction, for example, we would see earlier response times for when the sequence "stops being an [object]" than when it "starts being a [scene]." While we can estimate hysteresis in several ways, the most conservative estimate is the difference in response before any periods of uncertainty occur, such that the sequence is classified *with certainty* as an object or a scene depending on the presentation direction. We measured this by comparing responses to the prompt "stops being a ____" in the two presentation directions. If the hysteresis effect that we have observed in the preceding experiments is the byproduct of uncertainty in the decision, then there should not be a difference in these responses.

Of course, it may be that asking an observer to consider both objects and scenes in a single experiment has implications for the hysteresis effect. To address this possibility, we modified the task in Experiment 4B to only probe scenes; the experimenter never said the word *object* in the task description, and all observers were surveyed afterward to ensure that they did not consider objects in their response. This manipulation allowed us to evaluate the potential contributions of two-choice categorization on hysteresis. The results are relevant for evaluating whether the effect we report is likely to be visual or decisional. If our hysteresis effect is related to the hysteresis reported for low-level visual features, then it should be present in the detection of a single feature (as in Kleinschmidt et al., 2002, for example). However, if a simple change in the decisional task affects the hysteresis, we may consider the effect to be cognitive rather than perceptual.

Method

Twenty-four new observers (ages 19–29 years; 11 males and 13 females) were recruited from the Vanderbilt University study pool to participate in this Experiment 4A. Another 11 observers (ages 18–21 years; five males and six females) from the same population participated in Experiment 4B. All participants received either paid compensation or course credit for this study.

Stimuli, presentation, and experimental set-up of Experiment 4A were identical to those of Experiment 1, and the image sequences were presented as smooth movies at 15 frames/ second. Observers completed three total blocks of the classification task. In the first two blocks, they were asked to categorize each movie as "object" or "not object' or as "scene" or "not scene," rather than report a "change" between object and scene. In one of these single-category blocks, observers were asked to report when the movie "stopped being a ____," probing the period for which the transition was classified as the specific starting object or scene. In the other block, observers were asked to report when the movie "started being a ____," probing the period for which the movie was classified as the ending object or scene. The order of these blocks was counterbalanced across observers, and trial order was randomized as in Experiment 1. The third block, which appeared last for all participants, required the same classification as in Experiment 1 ("Report when the movie changed from ____ to ____"), providing a baseline measure of classification.

Experiment 4B also closely followed the design of Experiment 1. Observers completed four blocks of the scene detection task, classifying each transition as scene/not scene. Each trial asked when the image "stopped being a ____" or "started being a ____" based on the presentation direction and provided the corresponding scene label. Movie direction and presentation order were randomized and counterbalanced exactly as in Experiment 1; thus, trials asking the observer to report whether the images started or stopped being a scene were intermixed within each block. Afterward, observers answered a series of questions about their task performance. These questions scaled in specificity, first asking observers if they thought about *anything* other than scenes when making their responses and eventually asking if they made a specific categorization between scenes and objects (akin to the task in Experiment 1).

Results and Discussion

In Figure 7, we combined the responses from both single-category blocks to quantify uncertainty in the classification of the object/scene transitions. This analysis indicated that uncertainty was present both in the object-to-scene, $t(23) = 10.68$, $p < .00001$, Cohen's $d = 2.18$, and scene-to-object, $t(23) = 5.29$, $p < .00003$, Cohen's $d = 1.08$, presentation directions. The average duration of this uncertainty period was 1.01 s, far less than the duration of the hysteresis effect found in our previous experiments.

Moreover, it is apparent that hysteresis remained present even when we consider the uncertainty period: there is a section of the transition, highlighted in Figure 7, that is classified with certainty as either a scene or an object depending on which direction the movie is viewed. The difference between subjects' responses in the "stopped being a ____" condition is our most conservative estimate of the
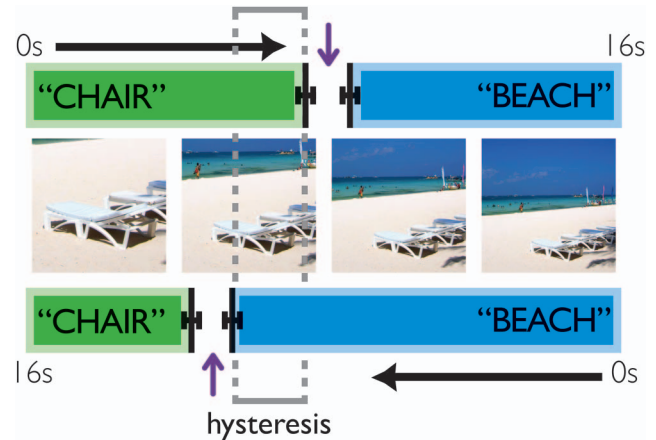

Figure 7. Results of Experiment 4 show uncertainty in the decision. The data from both single-category blocks, in which observers were asked to report when the sequence "stops being a ____" or "starts being a ____" closely resembles the predictions of the combined model in Figure 6. A period of uncertainty is present in the object/scene classification (marked by arrows), as is a distinct period of hysteresis (2.26 s). See online article for color version of this figure. Images courtesy of Asisbiz.

hysteresis effect, as it does not require the observer to accumulate any evidence that the transition has changed to the ending category. This difference is still 2.26 s, which is significantly greater than zero, one-sided $t(23) = 4.52$, $p < .0001$, Cohen's $d = 0.922$. Thus, even when we consider decisional uncertainty in our object/scene transitions, allowing a period to be called neither an object nor a scene, the hysteresis effect is not wholly eliminated.

We also compared the responses from the baseline classification block to those in the single-category blocks and found that the baseline category boundary of "changed from ___ to ____" fell within the window of uncertainty. There is no significant difference between the baseline responses and the average of the single-category responses in either presentation direction, $ts(23) < 1.11$, $ps > .280$, Cohen's $ds < .226$. Moreover, if we compare the averages of the single-category responses in the two presentation directions, we see a hysteresis effect of 3.32 s ($SE = 0.388$), which is comparable to our previous estimates of the hysteresis effect. This affirms that the single-category block reports were not qualitatively different from those made in the original object/scene categorization task.

In summary, Experiment 4A demonstrates that some uncertainty is present in the categorization of object/scene transitions, as evidenced by periods in which the movie was not classified as either an object or a scene in our single-category task blocks. However, even when this potential source of uncertainty is taken into account, a discernable hysteresis effect is still present in observers' judgments.

In Experiment 4B, we sought to quantify hysteresis for objects and scenes using a task in which observers were asked only to consider scenes in their reports. While Experiment 4A provided some evidence that hysteresis remains even when only one stimulus is probed, the design still required observers to consider both objects and scenes on subsequent trials within a block. Experiment 4B carefully limited the task to the singular detection of a scene: all instructions, from the initial study sign-up to the on-screen reminders, never mentioned objects. After the experiment, observ-

ers were asked about their attention to nonscene attributes of the image, including whether they consciously considered the object in each image when making their reports. One observer's data were excluded from analysis based on his answers.

For the remaining 10 observers, the hysteresis in scene-only reports was 3.10 s ($SE = 0.550$). This is statistically equivalent to the 3.07-s hysteresis found in Experiment 1, $t(9) = 0.076$, $p = .941$, Cohen's $d = 0.023$. With this control, we have demonstrated that hysteresis is not simply a consequence of considering two categorical interpretations of a stimulus. When subjects focus on a singular perceptual decision—whether they perceive a scene—the effect is equally robust.

Experiment 5: Providing Monetary Incentives to Reduce Hysteresis

In this experiment, we asked whether the hysteresis effect would still be found if observers were encouraged to respond as consistently as possible when performing the object/scene classification task. We hypothesized that as the hysteresis effect is quite pervasive, it would persist if observers remain focused on the object/scene discrimination. To test this hypothesis, we provided observers with monetary incentive to reduce the hysteresis effect in their responses, as well as feedback to facilitate learning and the use of strategies. We predicted that by the end of the experiment, observers would acquire and adopt cognitive strategies that would maximize consistency (and the monetary reward), independent of the instructed classification task. Critically, we were interested in the prevalence of hysteresis in the early blocks of this experiment, when observers were highly motivated to be consistent in their object/scene classification responses but did not have sufficient learning opportunity to develop specific strategies to do so.

Method

Thirty new observers from the Vanderbilt and Nashville community (ages 18–31 years; six males and 24 females) participated in this experiment for monetary compensation. The Vanderbilt University Institutional Review Board approved the study, including the use of performance-based payment.

The stimuli and task structure were identical to those of Experiment 1, in which observers watched movie-style image sequences and reported a single subjective boundary between object and scene with a key press. Observers completed five blocks of the category judgment task, viewing all 16 movies once in each direction in each block. Movies and directions were randomly ordered within the block, but the same movie could not appear twice in succession.

In the key manipulation of the current experiment, observers were instructed to perform the object/scene classification task with the goal of trying to make their responses as consistent as possible across the two presentation directions of each movie within each experimental block. Since each movie was presented once in each direction per block, measuring consistency within a block allowed us to provide clearly interpretable feedback to the observer about his or her performance. This feedback was given at the end of each block, which specified the number of points earned in that block as well as cumulatively. Consistency was defined as a temporal difference of less than 1 s, and participants were told that they

would earn a "bonus point" for every consistent response. In effect, we encouraged participants to eliminate the hysteresis difference between the two movie directions. Up to 16 bonus points could be earned in each block, and observers were told the number of bonus points they had earned after each block. We paid observers $0.20 for each bonus point earned, capped at $6 total bonus in possible payment.

The experiment lasted approximately 50 min. After completing the task, observers were asked 10 questions (either verbally or in writing) about their task performance and use of cognitive strategies that were relevant to or deviated from the primary classification task. The survey assessed use of strategy, motivation to earn bonus points, and attention to the object/scene classification task. It also asked participants if these variables changed throughout the experiment.

Three observers were excluded from analysis due to failure to perform the task adequately, as they had each earned less than 5 points total across the five experimental blocks. In contrast, the remaining observers earned an average of 45.33 points ($SE = 4.22$). We postulated that the outlier participants either failed to comprehend the task instructions or disregarded them.

Results

The results of Experiment 5 are plotted in Figure 8. As expected, observers demonstrated a steady increase in the number of points earned between each pair of successive blocks, all ts(26) > 2.35, ps $< .03$, Cohen's ds > 0.453. Participants earned an average of 4.5 points in Block 1 and 12.4 out of a total of 16 possible points in Block 5. These results indicated that as the experiment progressed, observers learned to perform more and more consistently at reporting a common time point for the object/scene boundary across the two movie directions. Accordingly, the average hysteresis difference across the five blocks decreased steadily: 3.14 s, 2.35 s, 1.60 s, 1.32 s, and 0.948 s in the last block; the decrease in hysteresis difference showed a significant main effect of block, $F(4, 104) = 17.555$, $p < .001$, partial $\eta^2 = .403$, and strong linear trend, $F(1, 26) = 30.87$, $p < .001$, partial $\eta^2 = .156$, as well as a smaller quadratic trend, $F(1, 26) = 4.77$, $p < .04$, partial $\eta^2 = .156$. The hysteresis effect is plotted by block number in Figure 8B, along with the average hysteresis effect by block for Experiment 1. Experiment 1 was identical to the current experiment save for the bonus point manipulation and showed no effect of block number on the size of the hysteresis effect.

In Experiment 5, observers exhibited a strong hysteresis effect (3.14 s) in the first block of the task, comparable to that in the first experiment, but these participants showed evidence of gradual learning such that they were able to nearly eliminate a difference in their responses by the last block (0.948 s). This decrease in the hysteresis effect was not present in Experiment 1 and can thus be attributed to the financial incentive and performance-based feedback given in the current experiment. So, how exactly did the observers in the current experiment learn to make increasingly consistent responses?

Twenty-three of 27 observers completed a 10-question survey about their performance and use of strategies; the remaining four observers spoke informally with the experimenter and reported strategies in line with the survey data. All surveyed observers ($N = 23$) reported using a cognitive strategy to perform the task, and
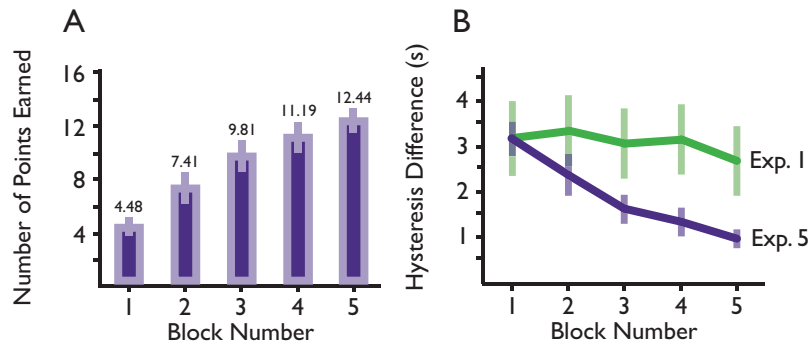
Figure 8.   Results of Experiment 5 are shown: providing monetary incentives to reduce hysteresis. Panel A: The number of points earned by observers across the five blocks of the experiment, out of a possible 16 points/block. Performance showed large improvement over the course of the experiment, confirming that observers developed strategies to make consistent responses. Panel B: The average length of the hysteresis difference of each block of the experiment, plotted alongside the same measure in Experiment 1. Experiment 1 was identical to the current design, save for the reward and feedback for consistency. In the first block of Experiment 5, observers exhibited a 3.14-s effect of hysteresis while performing the object/scene classification task; by the fifth block, most reported paying "little or no attention" to the object/scene classification task and had reduced their hysteresis difference to .948 s to meet the consistency demands. See online article for color version of this figure.

95.7% reported that their primary strategic goal was to earn bonus points. All observers reported using the visibility or alignment of specific objects to the frame of the viewed image; two observers (8.7%) also reported using counting to try to increase response consistency. No other specific strategies were reported. We also asked observers about their primary goals when performing the classification task over the course of the experiment; 82.6% reported that their cognitive strategy for the task changed over the experimental session, and 69.6% affirmed that after a point, they paid "little or no attention" to the original object/scene classification task. Of the 12 observers who reported remembering when they began ignoring the object/scene classification task, most reported starting from the second or third block ($M = 2.91$, $SE = 0.334$). Only one observer reported that this happened within the first block.

Thus, it appears that a sizeable majority of the observers used adaptive strategies that deviated from the original object/scene classification task to increase the consistency of their responses for performance-based earnings. The most commonly used strategy, reported by 100% of the surveyed observers, was attending to the alignment or appearance of specific objects in the frame. Although our survey did not probe in detail the time course of strategy adoption and classification task attention for each participant, we can venture that in the first block of the task, most observers paid some attention to the original classification task, but by the fifth block, observers were primarily relying on alternative cognitive strategies to maximize response consistency and ignoring the object/scene classification task. In the first block, the group-average hysteresis difference was 3.14 s, which was statistically comparable to the 3.07-s effect found in Experiment 1, $F(1, 35) = .005$, $p = .943$, partial $\eta^2 < .001$. In the fifth block, this difference was on average 0.948 s, below the task definition for consistency. The most consistent observer showed an average difference of just 145 ms in the fifth block of the experiment, not only responding at very similar points in two directions of the same movie but also performing this task across 16 distinct pairs presented in random-

ized order. We conclude that when observers are genuinely focused on performing the object/scene classification task in the early blocks, high motivation to provide consistent responses does not eliminate hysteresis. However, if observers employ an alternative response strategy based on local physical cues in the stimulus, rather than responding based on object/scene perception, they can successfully reduce the hysteresis difference in their behavioral response times.

Experiment 6: Specificity to Objects and Scenes

In our study, we have proposed that the perception of objects and scenes may be especially prone to hysteresis, as these stimuli frequently co-occur, yet have markedly different demands on processing. We explored the specificity of the hysteresis effect in this final experiment, quantifying the effect of stimulus category and of the similarity between particular exemplars. We created a set of novel image sequences that linearly ramped in contrast between two images. One of these images was always the scene end point (Frame 240) of the previously used movies. The other end point could be the corresponding object end point (Frame 1; congruent-object condition), a completely new object image (incongruent-object condition), or a face image. By selecting a new set of images for both the incongruent-object and face conditions, we ensured that familiarity with each nonscene image was matched between conditions. The task was identical to that in Experiment 1: observers reported when the sequence changed from the first image to the second; specific labels were provided for each object and scene, and "woman" or "man" was used for the faces.

Given that hysteresis for gradual changes in contrast has been demonstrated (Kleinschmidt et al., 2002), we expected some hysteresis to occur in each condition. However, we hypothesized that the magnitude of hysteresis in each of the three conditions would reflect our visual system's lifelong experiences with these categories of stimuli. Specifically, given the frequent co-occurrence of

objects and scenes, the magnitude of the effect would be greater for transitions between objects and scenes than between faces and scenes. Moreover, we predicted an additional increase in the magnitude of hysteresis for the congruent-object condition, since this best represents the kind of perceptual ambiguity between objects and scenes present in natural vision. Simply, we expect that the system underlying the effect described in Experiments 1–5 would be preferentially engaged by naturalistic pairings (a chair and office vs. a chair and face) and by those reflecting a coherent visual environment (a chair and office vs. a chair and lake).

Method

Fourteen new observers from the Vanderbilt undergraduate community (ages 18–21, six males and eight females) participated in this experiment for course credit.

For this experiment, we created sequences of contrast-ramping blends between two images. There were three conditions of image pairings: congruent-object, incongruent-object, and face. In each of these three conditions, the scene end point (Frame 240) of one of the previously used sequences was paired with a nonscene image. In the congruent-object condition, this scene was paired with the object end point (Frame 1) of the same sequence. In the incongruent-object condition and in the face condition, the scene end point images were paired with newly found object and face images. These new images were selected and cropped like the objects in the original sequences: each included an immediate scene background, but the object or face encompassed the majority of the frame. Moreover, as in the original sequences, there was considerable variability in the viewing angle of the object or face and in the perceptual depth and clutter of the scene background; for example, some face images included shoulders, hats, or even part of another face. The new object stimuli consisted of a bedside table, bike, car, chair, coffee maker, computer, firepit, soccer goal, hat, houseplant, lamp, pool float, shopping cart, table, and tray; the face images were eight male and eight female faces. All images were 500×500 pixels, grayscale, and luminance histograms were matched across the full set of images using the SHINE toolbox for Matlab (Willenbockel et al., 2010).

As described, the congruent-object condition paired the corresponding scene and object end points of the sequences used in Experiments 1–5. For the incongruent-object condition, a randomized pairing of the scene and one of the 16 new object images was determined for each observer and repeated throughout the experiment, so that familiarity with each of the pairs was equal. The same randomized pairing was made for the face condition. On each trial, the two images complementarily ramped in contrast in 240 linear steps, so that the trial began with 100% of Image 1 and ended with 100% of Image 2. The images were presented at a rate of 15 frames/second, and each trial was 16 s long. Each pair in the three conditions (congruent-object, incongruent-object, and face) was presented in both directions in each block, for a total of 96 trials per block. Each observer completed two blocks of the task, and the direction in which each blend was first presented was counterbalanced across observers.

As in the previous experiments, observers were instructed to "Press a key when the movie changes from" one stimulus to the other. The entire experiment lasted approximately 60 min.

Results and Discussion

The results of Experiment 6 are shown in Figure 9. Hysteresis was found in each of the three conditions: 4.81 s for same-sequence object/scene pairs ($SE = 0.521$ s), 4.14 s for novel object/scene pairs ($SE = 0.512$ s), and 3.41 s for face/scene pairs ($SE = 0.419$ s). A repeated-measures ANOVA revealed a significant main effect of condition, $F(2, 26) = 7.683$, $p < .003$, partial $\eta^2 = .371$, of movie direction, $F(1, 13) = 15.895$, $p < .003$, partial $\eta^2 = .550$, and an interaction effect between these two factors, $F(2, 26) = 19.193$, $p < .001$, partial $\eta^2 = .596$. The effect of block was not significant, nor did it interact with the other effects ($Fs < 1.19$, $ps > .322$, partial $\eta^2 < .083$). We performed paired t tests to directly compare the magnitude of hysteresis across conditions and found significant differences among all three: hysteresis was larger for incongruent object/scene pairs than for face/scene, $t(13) = 3.50$, $p < .004$, Cohen's $d = 0.936$, and even larger for congruent object/scene pairs than for those with incongruent objects/scenes, $t(13) = 3.19$, $p < .008$, Cohen's $d = 0.854$.

From these results, we can see that hysteresis is not equal for all categories or even all exemplars of a certain category. Not only does hysteresis occur more robustly between objects and scenes than between faces and scenes but also its magnitude appears to be mediated by the congruence of the particular objects and scenes used. This is consistent with the idea that hysteresis arises from our common experience of viewing objects and scenes in a single environment. Natural vision more frequently requires us to switch between object and scene perception within a single environment (chair to beach) than between incongruous stimuli (coffeemaker to beach); the efficiency of these "congruent" switches can be informed by our perceptual memory of that environment, whereas those between pairs that lack continuity cannot.

While one could argue that the differences found in these artificial sequences are driven by low-level dissimilarities, it is critical to note that these differences are inherent to natural perception of high-level categories; in this way, the statistical regularities present are not confounds but meaningful co-occurrences. Differential hysteresis for such regularities is far beyond the predictions of existing literature in low-level vision (Buckthought et al., 2008; Fender & Julesz, 1967; Hock et al., 1994; 2005; Julesz, 1974; Kleinschmidt et al., 2002; Williams & Sekuler, 1986). Future work should clarify where along the processing hierarchy this high-level hysteresis effect first arises.

General Discussion

In this series of experiments, we have demonstrated a persistent effect of hysteresis in the categorization of objects and scenes when observers view dynamic sequences of images that transition between two canonical views. That is, an image sequence is judged to be an object for a much longer time period when the transition starts at the object and moves toward the scene than when it is presented in the opposite direction. Throughout the study, we compared responses between presentation directions of the same movie, showing it from object-to-scene and from scene-to-object. This allowed us to equate the physical properties of the stimulus and directly measure the effect of viewing history. This effect was on the order of 2–3 s for our 16-s movie sequences and was
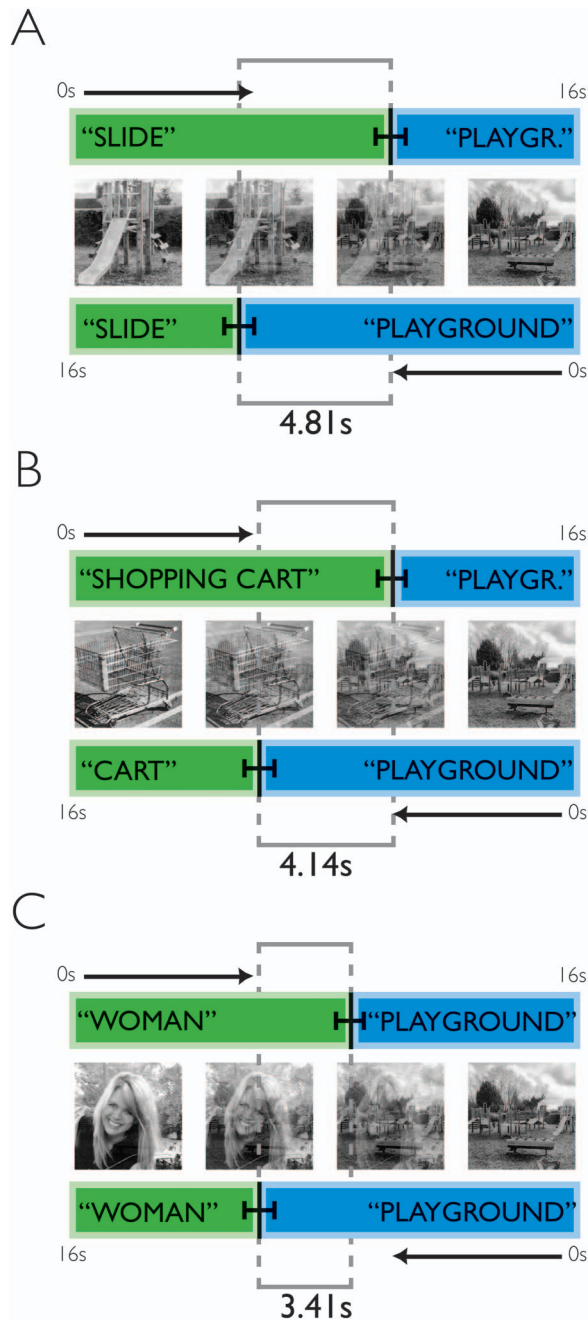
Figure 9. Results of Experiment 6 are shown: specificity to objects and scenes. Average length of hysteresis difference for the (Panel A) congruent-object condition, (Panel B) incongruent-object condition, and (Panel C) face condition. In each condition, the average results are displayed on a graphic depiction of a single trial for clarity. We found significant differences in the magnitude of hysteresis in each of the conditions, suggesting a specificity of the effect to objects and scenes, particularly when they depict a single visual environment. See online article for color version of this figure. Playground images: © Tom Morris. Published under a Creative Commons Attribution license. Shopping cart images: © Stilfehler. Published under a Creative Commons Attribution license. Due to permissions considerations, the image of one of the authors (Sonia Poltoratski) appears here representing one of the study stimuli.

persistent across a widely variable set of images that contained different amounts of clutter and different types of object content.

In the first four experiments, we quantified the contributions of properties of the stimulus and of the instructed decision on the hysteresis effect. By presenting the sequences as series of static images (Experiment 2), we found that the effect does not require motion cues. Next, we established that unidirectional progression through the image sequence is necessary for hysteresis, while passive viewing of the stimuli is not (Experiments 3A and 3B). This suggests that a degree of perceptual continuity is necessary, which is not unusual given the regularities and perceptual continuity present in our viewing of objects and scenes in the real world. In Experiment 4A, we asked whether the presence of uncertainty in the object/scene decision could fully account for the pattern of responses that we had attributed to hysteresis. Even though our manipulation found a discernable period of uncertainty, this did not fully account for the hysteresis effect; there was still a period of the sequence for which it is classified with certainty as an object or a scene depending on the prior viewing history. Moreover, the effect was not eliminated when observers were instructed to consider only scenes in their responses, rather than to make a binary classification (Experiment 4B).

Next, we sought to test whether the hysteresis effect would be found if observers were given monetary incentive and feedback to respond as consistently as possible when performing the object/scene classification task (Experiment 5). We found that observers cannot willfully prevent hysteresis from influencing their categorization of dynamic objects and scenes when they are genuinely engaged in performing this type of judgment; however, observers can learn adaptive strategies to make consistent responses on our dynamic, naturalistic stimulus sequences. Finally, we sought to directly evaluate the specificity of hysteresis to objects and scenes, particularly pairs taken from a single environment (Experiment 6). Using a linear contrast-ramp procedure, we found that hysteresis was significantly greater for object/scene pairs of images than for face/scene pairs, with an additional increase in hysteresis for congruent object/scene pairs over random pairings.

These six experiments demonstrate a persistent effect of hysteresis in categorizing naturalistic sequences that present changing views of objects and scenes. The dynamic stimuli that we used were designed to approximate the way in which objects and scenes are encountered in natural vision, changing from canonical views of each category by zooming and panning. The finding of hysteresis, particularly its relative specificity to the perception of objects and scenes, implies a unified, cooperative framework that encompasses key elements of our natural visual experience of these two types of stimuli.

Our natural vision seems unified and seamless, even though the input that our visual system receives is not. Even within the visual modality, features such as form, motion, color, and depth are registered and processed by quite distinct pathways or channels in the visual system (Livingstone & Hubel, 1987, 1988), such that they are not automatically combined across saccades unless attention is deployed to form detailed representations (Blackmore, Brelstaff, Nelson, & Troscianko, 1995; Henderson, 1997; Hollingworth & Henderson, 2002). How does the visual system integrate such fragmented input, not only within a single object or view but also across dynamic changes

in viewpoint during natural vision? The mechanism of hysteresis, which has been used to describe the behavior of other biological and mechanical systems that require transitions between distinct states, can provide the visual system with a way to reduce ambiguity at the boundary between two percepts. Using naturalistic movie stimuli and a simple classification task, we have identified a period of perceptual ambiguity likely present in the natural viewing of objects and scenes and suggest that the visual system employs the mechanism of hysteresis to resolve this ambiguity. In doing so, we have extended the application of hysteresis to the visual system beyond simple features (Buckthought et al., 2008; Fender & Julesz, 1967; Hock et al., 1993; 2005; Julesz, 1974; Kleinschmidt et al., 2002; Williams & Sekuler, 1986) or the identification of exemplars within a particular stimulus type, such as line-drawn object shapes or facial emotion (Sacharin et al., 2012; You et al., 2011).

This sort of processing system is especially well adapted for objects and scenes, particularly as they occur in natural vision: separate, but together. Object perception is typically localized in the lateral occipital complex (Grill-Spector, Kourtzi, & Kanwisher, 2001), while scene perception has been linked to a cortical network including parahippocampal gyrus, the retrosplenial cortex, and the transverse occipital sulcus (Dilks et al., 2013;. Epstein, 2008; Epstein & Kanwisher, 1998; Park & Chun, 2009). Models of scene and object perception that consider both in unison have differentiated them by scale, describing the "fine edges" of objects or "course blobs" of scenes that carry the majority of information about their identity (Greene & Oliva, 2009a; Kersten et al., 2004; Oliva & Schyns, 1997; Oliva & Torralba, 2001; Renninger & Malik, 2004; Torralba & Oliva, 2002). Scene and object perception also place considerable demands on the limited attentional resources of the visual system. Hollingworth and Henderson (2002) have emphasized the transience of object representations by showing that directed attention is needed to bind features across saccades. Similarly, recent research has dispelled the idea that extracting even basic gist information from natural scenes requires no attention (Cohen et al., 2011; Rousselet et al., 2004).

Although object processing and scene processing are largely separable, we always encounter objects in a scene, and most scenes contain at least one object. In these situations, it follows that the visual system should sometimes prioritize either object or scene perception to maximize the distinct processing of both states, "switching" based on both attentional and scale cues. Hysteresis can maximize the efficiency of switching between object perception and scene perception by exploiting the rich perceptual continuity of natural vision. Moreover, unnecessary switches can be reduced if the system considers recent perceptual history, which is likely to be informative within a single environment. As the visual system is undoubtedly shaped by the visuotemporal continuities we encounter, a system for efficiently and continuously processing objects and scenes seems more likely than one for a pair of stimuli with a lesser functional relationship. In support of this, we find that hysteresis is greater for object/scene pairs than for scene/face pairs, and even greater when the object/scene pair is taken from a single environment (Experiment 6). An intriguing future direction would be to explore hysteresis in the perception of bodies and faces, which similarly co-occur and differ mostly as a function of spatial scale.

The mechanism of hysteresis has previously been modeled as a neural network whose nonlinear excitatory and inhibitory properties can give rise to "coalitions" between similarly tuned elements (Williams & Sekuler, 1986). These dynamic interactions can serve to enhance the signal-to-noise ratio of the system and allow a decision between two states to be made during periods of rapid change in the external stimulus (You et al., 2011). This interactive system is reminiscent of neural models of binocular rivalry, which have incorporated within-state co-excitation and between-state inhibition to account for the temporal dynamics of bistable perception (Laing & Chow, 2002). Generally, hysteresis is a means for a single system with two end points to efficiently function in intermediate, noisy states. The current findings imply such a cooperative, single-system relationship between the perception of objects and scenes; that is, their perceptions are systematically linked in a way that can be both cooperative and competitive. Some interactive benefits in processing have been reported when scenes and objects provide consistent diagnostic cues (Biederman, 1972; Palmer, 1975; for a review, see Bar, 2004; Davenport, 2007; Davenport & Potter, 2004; MacEvoy & Epstien, 2011; Oliva & Torralba, 2007). The competitive aspect of object/scene perception, while implied by their cognitive demands (Cohen et al., 2011; Rousselet et al., 2004; VanRullen & Koch, 2003) and scale-based separability (Kersten et al., 2004; Navon, 1977; Oliva & Torralba, 2001) deserves further study.

One possibility is that the effects of hysteresis observed here may partly arise from visual cooperation that occurs within brain regions involved in scene processing and those involved in object processing, and competition between these regions. Our current results have demonstrated that hysteresis occurs not only in the perception of features that can be represented by a local pool of neurons but also between objects and scenes, which have typically been localized in distinct neural regions (Epstein & Higgins, 2007; Epstein & Kanwisher, 1998; Grill-Spector et al., 2001). However, we hypothesize that the co-occurrence of these stimuli in natural vision necessitates neural connectivity. The presence of large-scale connections between object- and scene-selective cortical areas has been suggested by a recent study that employed transcranial magnetic stimulation (TMS); Mullin and Steeves (2011) found that when TMS was used to temporarily disrupt processing in left lateral occipital complex, observers became impaired at classifying objects but actually improved in their performance at classifying scenes. However, in additional work, researchers must ask how engaging object perception may interfere with scene perception, and vice versa, when they compose a single visual environment.

Object processing and scene processing have largely been studied with simpler, static images. Although the naturalistic image sequences used in Experiments 1–5 were quite minimal and somewhat unnaturally smooth, the study set-up afforded observers complete freedom in eye and head movements. In this way, our paradigm allowed for a degree of discontinuity and interobserver variability. Thus, although Experiment 3 showed that some perceptual continuity was necessary for the effect to occur, this continuity need not be perfect; Experiment 2, in which the sequences were presented with 300-ms interleaved blank periods, speaks further to this. However, in future work, researchers should quantify the timing of the effect and how easily it is disrupted by inconsistencies or abrupt shifts in viewpoint. Similarly, shifts of attention between objects and scenes—independent of changes in viewpoint—may exhibit their own hysteresis effects. In fact, visual search patterns have been found to show considerable

interobserver consistency, suggesting a role of prior experience (Hidalgo-Sotelo & Oliva, 2010).

While the current study has explored many of the possible perceptual and decisional contributions to hysteresis, the novelty of this finding begs considerable further research. Attention, which our measures did not directly probe, vastly impacts the building of representations of both objects (Hollingworth & Henderson, 2002) and scenes (Rensink et al., 1997) across saccades and similar discontinuities. Attention undoubtedly plays a complex role in the unified, interactive system of object/scene perception we have described. While we tested the effect of different prompts and labels in Experiments 4A and 4B, we did not have a direct measure of the observers' locus of attention. However, we might presume that the instruction to respond based on a single category or two categories would encourage observers to consider different sets of features in their viewing of the sequences. That this manipulation yielded consistent measures of hysteresis may imply that attention does not mediate the effect itself, although attention is needed to perform the object/scene classification task (Expt. 6). This is consistent with our interpretation of hysteresis as a high-level perceptual bias, but should be quantified in additional research.

In summary, we have considered objects and scenes as they occur in dynamic natural vision and found a consistent effect of perceptual history on their classification. We have identified this effect as hysteresis and suggest that it allows for the seamless and efficient switching between two states of perception. Its somewhat specific presence in the perception of objects and scenes implies a unified, cooperative system for their processing, which is likely informed by the co-occurrence of these stimuli in natural vision.

References

Angeli, D., Ferrell, J. E., & Sontag, E. D. (2004). Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *PNAS: Proceedings of the National Academy of Sciences of the United States of America, 101,* 1822–1827. doi:10.1073/pnas.0308265100

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5,* 617–629. doi:10.1038/nrn1476

Bartels, A. (2009). Visual perception: Converging mechanisms of attention, binding, and segmentation?. *Current Biology, 19,* R300–R302. doi:10.1016/j.cub.2009.02.014

Bertotti, G. (1998). *Hysteresis in magnetism: For physicists, materials scientists, and engineers.* Chestnut Hill, MA: Academic Press.

Biederman, I. (1972, July 7). Perceiving real-world scenes. *Science, 177,* 77–80. doi:10.1126/science.177.4043.77

Biederman, I., & Gerhardstein, P. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 1162–1182. doi:10.1037/0096-1523.19.6.1162

Blackmore, S. J., Brelstaff, G., Nelson, K., & Troscianko, T. (1995). Is the richness of our visual world an illusion? Transsaccadic memory for complex scenes. *Perception, 24,* 1075–1081. doi:10.1068/p241075

Blanchard, O. J., & Summers, L. H. (1986). Hysteresis and the European unemployment problem. *NBER Macroeconomics Annual, 1,* 15–90. doi:10.2307/3585159

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436. doi:10.1163/156856897X00357

Brokate, M., & Friedman, A. (1989). Optimal design for heat conduction problems with hysteresis. *SIAM Journal on Control and Optimization, 27,* 697–717. doi:10.1137/0327037

Buckthought, A., Kim, J., & Wilson, H. R. (2008). Hysteresis effects in stereopsis and binocular rivalry. *Vision Research, 48,* 819–830. doi:10.1016/j.visres.2007.12.013

Bülthoff, H., Edelman, S., & Tarr, M. (1995). How are three-dimensional objects represented in the brain? *Cerebral Cortex, 5,* 247–260. doi:10.1093/cercor/5.3.247

Burr, D. C., & Ross, J. (1986). Visual processing of motion. *Trends in Neurosciences, 9,* 304–307. doi:10.1016/0166-2236(86)90088-3

Burr, D. C., Ross, J., & Morrone, M. (1986). Smooth and sampled motion. *Vision Research, 26,* 643–652. doi:10.1016/0042-6989(86)90012-X

Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological Science 22,* 1165–1172. doi:10.1177/0956797611419168

Collin, C. A., & Mcmullen, P. A. (2005). Subordinate-level categorization relies on high spatial frequencies to a greater degree than basic-level categorization. *Perception & Psychophysics, 67,* 354–364. doi:10.3758/BF03206498

Davenport, J. L. (2007). Consistency effects between objects in scenes. *Memory & Cognition, 35,* 393–401. doi:10.3758/BF03193280

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science, 15,* 559–564. doi:10.1111/j.0956-7976.2004.00719.x

De Cesarei, A., & Loftus, G. R. (2011). Global and local vision in natural scene identification *Psychonomic Bulletin & Review, 18,* 840–847. doi:10.3758/s13423-011-0133-6

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron, 73,* 415–434. doi:10.1016/j.neuron.2012.01.010

Dilks, D. D., Julian, J. B., Paunov, A. M., & Kanwisher, N. (2013). The occipital place area is causally and selectively involved in scene perception. *Journal of Neuroscience, 33,* 1331–1336. doi:10.1523/JNEUROSCI.4081-12.2013

Epstein, R. A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences, 12,* 388–396. doi:10.1016/j.tics.2008.07.004

Epstein, R. A., & Higgins, J. S. (2007). Differential parahippocampal and retrosplenial involvement in three types of visual scene recognition. *Cerebral Cortex, 17,* 1680–1693. doi:10.1093/cercor/bhl079

Epstein, R., & Kanwisher, N. (1998, April 9). A cortical representation of the local visual environment [Letter to the editor]. *Nature, 392,* 598–601. doi:10.1038/33402

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance, 31,* 1476. doi:10.1037/0096-1523.31.6.1476

Fei-Fei, L., & Perona, P. (2005). A Bayesian hierarchical model for learning natural scene categories. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2,* 524–531.

Fei-Fei, L., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences, 99,* 9596–9601.

Fender, D., & Julesz, B. (1967). Extension of Panum's fusional area in binocularly stabilized vision. *Journal of the Optical Society of America, 57,* 819–830. doi:10.1364/JOSA.57.000819

Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Cognitive Brain Research, 16,* 123–144. doi:10.1016/S0926-6410(02)00244-6

Goldstone, R. L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition, 52,* 125–157. doi:10.1016/0010-0277(94)90065-5

Greene, M. R., & Oliva, A. (2009a). The briefest of glances: The time course of natural scene understanding. *Psychological Science, 20,* 464–472. doi:10.1111/j.1467-9280.2009.02316.x

Greene, M. R., & Oliva, A. (2009b). Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cognitive Psychology 58,* 137–176. doi:10.1016/j.cogpsych.2008.06.001

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision research, 41,* 1409–1422.

Grill-Spector, K., Kushnir, T., Edelman, S., & Avidan, G. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron, 24,* 187–203. doi:10.1016/S0896-6273(00)80832-6

Henderson, J. M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science, 8,* 51–55. doi:10.1111/j.1467-9280.1997.tb00543.x

Hidalgo-Sotelo, B., & Oliva, A. (2010). Person, place, and past influence eye movements during visual search. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 820–825). Austin, TX: Cognitive Science Society.

Hock, H. S., Bukowski, L., Nichols, D. F., Huisman, A., & Rivera, M. (2005). Dynamical vs. judgmental comparison: Hysteresis effects in motion perception. *Spatial Vision, 18,* 317–335. doi:10.1163/1568568054089393

Hock, H. S., Kelso, J. A. S., & Schöner, G. (1993). Bistability and hysteresis in the organization of apparent motion patterns. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 63–80. doi:10.1037/0096-1523.19.1.63

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 113. doi:10.1037/0096-1523.28.1.113

Jiles, D. C., & Atherton, D. L. (1986). Theory of ferromagnetic hysteresis. *Journal of Magnetism and Magnetic Materials, 61,* 48–60. doi:10.1016/0304-8853(86)90066-1

Julesz, B. (1974). Cooperative phenomena in binocular depth perception. *American Scientist, 62,* 32–43.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 55,* 271–304. doi:10.1146/annurev.psych.55.090902.142005

Kleinschmidt, A., Buchel, C., Hutton, C., Friston, K. J., & Frackowiak, R. S. J. (2002). The neural structures expressing perceptual hysteresis in visual letter recognition. *Neuron, 34,* 659–666. doi:10.1016/S0896-6273(02)00694-3

Kourtzi, Z., Erb, M., Grodd, W., & Bülthoff, H. H. (2003). Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cerebral Cortex, 13,* 911–920. doi:10.1093/cercor/13.9.911

Kourtzi, Z., & Kanwisher, N. (2001, August 24). Representation of perceived object shape by the human lateral occipital complex. *Science, 293,* 1506–1509. doi:10.1126/science.1061133

Laing, C. R., & Chow, C. C. (2002). A spiking neuron model for binocular rivalry. *Journal of Computational Neuroscience, 12,* 39–53. doi:10.1023/A:1014942129705

Leopold, D. A., Wilke, M., Maier, A., & Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. *Nature Neuroscience, 5,* 605–609. doi:10.1038/nn0602-851

Livingstone, M., & Hubel, D. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience, 7,* 3416–3468.

Livingstone, M., & Hubel, D. (1988, May 6). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science, 240,* 740–749. doi:10.1126/science.3283936

MacEvoy, S., & Epstein, R. (2011). Constructing scenes from objects in human occipitotemporal cortex. *Nature Neuroscience, 14,* 1323–1329. doi:10.1038/nn.2903

Maier, A., Wilke, M., Logothetis, N. K., & Leopold, D. A. (2003). Perception of temporally interleaved ambiguous patterns. *Current Biology, 13,* 1076–1085. doi:10.1016/S0960-9822(03)00414-7

Mullin, C. R., & Steeves, J. K. R. (2011). TMS to the lateral occipital cortex disrupts object processing but facilitates scene processing. *Journal of Cognitive Neuroscience, 23,* 4174–4184. doi:10.1162/jocn_a_00095

Nakayama, K., He, Z. J., & Shimojo, S. (1995). Visual surface representation: A critical link between lower level and higher level vision. In S. M. Kosslyn & D. N. Osherson (Eds.), *An invitation to cognitive science* (pp. 1–70). Cambridge, MA: MIT Press.

Nakayama, K., & Silverman, G. (1984). Temporal and spatial characteristics of the upper displacement limit for motion in random dots. *Vision Research, 24,* 293–299. doi:10.1016/0042-6989(84)90054-3

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353–383. doi:10.1016/0010-0285(77)90012-3

Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology, 34,* 72–107. doi:10.1006/cogp.1997.0667

Oliva, A., & Schyns, P. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology, 41,* 176–210. doi:10.1006/cogp.1999.0728

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision, 42,* 145–175. doi:10.1023/A:1011139631724

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences, 11,* 520–527. doi:10.1016/j.tics.2007.09.009

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3,* 519–526.

Palmeri, T. J., & Gauthier, I. (2004). Visual object understanding. *Nature Reviews Neuroscience, 5,* 291–303. doi:10.1038/nrn1364

Park, S., & Chun, M. M. (2009). Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *NeuroImage, 47,* 1747–1756. doi:10.1016/j.neuroimage.2009.04.058

Parker, D. M., Lishman, J. R., & Hughes, J. (1996). Role of coarse and fine information in face and object processing. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1448–1466. doi:10.1037/0096-1523.22.6.1448

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116,* 220–224. doi:10.1037/0033-2909.116.2.220

Pearson, J., & Brascamp, J. (2008). Sensory memory for ambiguous vision. *Trends in Cognitive Sciences, 12,* 334–341. doi:10.1016/j.tics.2008.05.006

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442. doi:10.1163/156856897X00366

Renninger, L. W., & Malik, J. (2004). When is scene identification just texture recognition? *Vision Research, 44,* 2301–2311. doi:10.1016/j.visres.2004.04.006

Rensink, R. A., O'Regan, J. K. O., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368–373. doi:10.1111/j.1467-9280.1997.tb00427.x

Riesenhuber, M., & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience, 3,* 1199–1204. doi:10.1038/81479

Rousselet, G. A., Thorpe, S. J., & Fabre-Thorpe, M. (2004). Processing of one, two, or four natural scenes in humans: The limits of parallelism. *Vision Research, 44,* 877–894. doi:10.1016/j.visres.2003.11.014

Sacharin, V., Sander, D., & Scherer, K. R. (2012): The perception of changing emotion expressions. *Cognition & Emotion, 26,* 1273–1300. doi:10.1080/02699931.2012.656583

Schmitt, O. H. (1938). A thermionic trigger. *Journal of Scientific Instruments, 15,* 24. doi:10.1088/0950-7671/15/1/305

Schyns, P. G. (1998). Diagnostic recognition: Task constraints, object information, and their interactions. *Cognition, 67,* 147–179. doi:10.1016/S0010-0277(98)00016-X

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science, 5,* 195–200. doi:10.1111/j.1467-9280.1994.tb00500.x

Shulman, G. L., & Wilson, J. (1987). Spatial frequency and selective attention to local and global information. *Perception, 16,* 89–101. doi:10.1068/p160089

Tarr, M., & Bülthoff, H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1494–1505. doi:10.1037/0096-1523.21.6.1494

Tarr, M., Williams, P., Hayward, W., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint-dependent. *Nature Neuroscience, 1,* 275–277. doi:10.1038/1089

Todd, J. T. (2004). The visual perception of 3D shape. *Trends in Cognitive Sciences, 8,* 115–121. doi:10.1016/j.tics.2004.01.006

Torralba, A., & Oliva, A. (2002). Depth estimation from image structure. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24,* 1226–1238. doi:10.1109/TPAMI.2002.1033214

VanRullen, R., & Koch, C. (2003). Competition and selection during visual processing of natural scenes and objects. *Journal of Vision, 3*(1), 8. doi:10.1167/3.1.8

Vogel, J., Schwaninger, A., Wallraven, C., & Bülthoff, H. H. (2006). Categorization of natural scenes: Local vs. global information. In R. W. Fleming & S. Kim (Eds.), *Proceedings of the 3rd symposium on applied perception in graphics and visualization, Boston, Massachusetts* (pp. 33–40). New York, NY: ACM Press.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods, 42,* 671–684. doi:10.3758/BRM.42.3.671

Williams, D., & Sekuler, G. B. R. (1986, November 20). Hysteresis in the perception of motion direction as evidence for neural cooperativity [Letter to the editor. *Nature, 324,* 253–255. doi:10.1038/324253a0

Wolfe, J. M., Oliva, A., Horowitz, T. S., Butcher, S. J., & Bompas, A. (2002). Segmentation of objects from backgrounds in visual search tasks. *Vision Research, 42,* 2985–3004. doi:10.1016/S0042-6989(02)00388-7

You, H., Meng, Y., Huan, D., & Wang, D.-H. (2011). The neural dynamics of hysteresis in visual perception. *Neurocomputing, 74,* 3502–3508. doi:10.1016/j.neucom.2011.06.004

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at http://notify.apa.org/ and you will be notified by e-mail when issues of interest to you become available!