Philosophical Psychology

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713441835

Art, Artists, and Perception: A Model for Premotor Contributions to Perceptual Analysis and Form Recognition
William Seeley; Aaron Kozbelt

Online Publication Date: 01 April 2008
To cite this Article: Seeley, William and Kozbelt, Aaron (2008) 'Art, Artists, and Perception: A Model for Premotor Contributions to Perceptual Analysis and Form Recognition', Philosophical Psychology, 21:2, 149 - 171
To link to this article: DOI: 10.1080/09515080801976573
URL: http://dx.doi.org/10.1080/09515080801976573

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf
This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Art, Artists, and Perception: A Model for Premotor Contributions to Perceptual Analysis and Form Recognition

William Seeley and Aaron Kozbelt

Artists, art critics, art historians, and cognitive psychologists have asserted that visual artists perceive the world differently than nonartists and that these perceptual abilities are the product of knowledge of techniques for working in an artistic medium. In support of these claims, Kozbelt (2001) found that artists outperform nonartists in visual analysis tasks and that these perceptual advantages are statistically correlated with drawing skill. We propose a model to explain these results that is derived from a diagnostic framework for object recognition and recent research in cognitive neuroscience on selective visual attention. This research demonstrates that endogenous shifts of visual attention enhance the encoding of expected features in the visual field and inhibit the perception of potential distracters. Moreover, it demonstrates complementary roles for spatial schemata and motor plans in visual attention. We argue that artists develop novel spatial schemata, which enable them to recognize and reproduce stimulus features sufficient for adequate artistic production in a medium, and that these schemata become encoded as motor plans as artists develop technical proficiency in a medium. We hypothesize that artists’ perceptual advantages can therefore, be explained by the role spatial schemata and motor plans play in selective attention.

Keywords: Attention; Drawing Skill; Neuroscience; Object Recognition; Perception; Premotor; Visual Art

William Seeley is a Visiting Assistant Professor in the Department of Philosophy Franklin and Marshall College. Aaron Kozbelt is an Associate Professor in the Department of Psychology Brooklyn College City University of New York. Correspondence to: William Seeley, Department of Philosophy, Box 3003, Franklin and Marshall College, Lancaster, PA, 17604-3003, USA. Email: william.seeley@fandm.edu

ISSN 0951-5089 (print)/ISSN 1465-394X (online)/08/020149-23 © 2008 Taylor & Francis
DOI: 10.1080/09515080801976573
1. Introduction

Visual artists are able to create drawings and paintings that appear quite realistic. In contrast, most nonartists have difficulty producing even moderately convincing renderings of objects and scenes (Cohen & Bennett, 1997). Superficially this performance difference is surprising. The shapes of objects and the formal structure of scenes appear to be transparent to ordinary perception. It would seem, therefore, that all one needs to do to accurately depict something is focus one’s attention and carefully copy what one perceives. Artists (e.g., Leonardo and Ingres; see Goldwater & Treves, 1972), art critics and historians (Fry, 1909/1981; Ruskin, 1857/1971), and cognitive psychologists (Cohen, 2005; Cohen & Bennett, 1997; Mitchell, Ropar, Ackroyd, & Rajendran, 2005) have argued that this common sense story of drawing skill is essentially correct. However, practical knowledge biases perception in ordinary contexts and causes agents to misperceive stimuli. Therefore, practical knowledge interferes with the ability to render the formal geometry of scenes and objects accurately. Artists’ methods, including drawing skill, are means to overcome these perceptual biases. Thus, artists, art critics, and psychologists have argued that artists are better at depiction because they, unlike untrained viewers, learn to perceive the world as if it were an artwork (e.g., as a stimulus to be drawn, painted, modeled, etc.).

The following claims are representative of this view of artists’ perceptual advantages (APA):

APA1. Artists are better at the type of visual analysis necessary for depiction in their medium.

APA2. These perceptual advantages are derived from technical proficiency with the tools of an artistic medium.

Empirical research by Kozbelt (2001) supports (APA1) and provides tentative evidence for (APA2). Kozbelt examined the performance of artists and nonartists on sets of drawing tasks and perception tasks. Artist participants were undergraduate art or design majors. Nonartists were undergraduates not majoring in art or design. In the drawing tasks, participants copied line drawings of familiar objects, letters, and abstract patterns and generated their own line drawings of familiar objects from life and from photographs. Instructions for the drawing tasks emphasized copying stimuli accurately; accuracy was also the criterion judges used to rate the drawings. In the perception tasks, participants were asked to identify the subjects of blurred photographs, find abstract geometric figures embedded in complex visual patterns,
and identify figures in gestalt completion tasks. Figure 1a shows a sample blurred photograph. It depicts several mourning women with their heads covered by scarves. To identify the subject of this photograph, the image must be segmented into objects, figure-ground and relative depth relations must be established, some information must be discarded as spurious, and the features and contours of the objects must be matched to category knowledge and integrated into a unified interpretation of the scene. Similar visual analyses are also necessary for accurate drawing: understanding the relative depth of objects, choosing relevant visual information for inclusion in the depiction, and emphasizing features and contours salient to object identity.

Artists outperformed nonartists in drawing tasks with an average difference of 0.70 SDs across tasks, a large effect size. Drawing experience was the critical variable

Figure 1. (a) Sample Blurred Photograph. The Photograph depicting several mourning women with their heads covered in heavy scarves (Kozbelt, 2001); (b) circles identify several T-junctions (locations where one contour is interrupted by, and seems to disappear behind, another) diagnostic for the relative depth and forms of objects in image 1a (Hamm, 1963, p. 48); (c) participants had to discard spurious feature conjunctions in order to locate a target shape within a complex pattern to solve this sample embedded figures task.
differentiating artist from nonartist participants. Therefore, this finding is not surprising. However, artists also outperformed nonartists in the perception tasks, with an average difference of 0.72 SDs, again a large effect. There is plausible overlap in the processing required to do both kinds of tasks. Nonetheless, the perception tasks had nothing directly to do with drawing. Therefore, this result is of particular interest.

Three aspects of Kozbelt’s study are germane to our discussion. First, each perception task required participants to recover features constitutive of the form of depicted objects from degraded images (Figure 1b and c). Therefore, consistent with (APA1), artists’ superior performance in the perception tasks demonstrates an advantage in visual analysis that confers an advantage in form recognition. Second, across all participants, performance on drawing and perception tasks was positively correlated, $r(44) = 0.63, p < 0.01$. Thus, if a participant was good at drawing, that participant also tended to do well on the perception tasks. Lastly, partial correlational analyses of these data revealed a relationship between drawing skill and perception. When perceptual advantages were statistically controlled for, artists outperformed nonartists on the drawing tasks by an average of 0.34 SDs, a small to medium effect size, $F(1,42) = 5.47, p < 0.05$. This result indicates that artists’ superior ability in visual analysis tasks was insufficient to explain their success at drawing. However, artists did not reliably outperform nonartists on the perception tasks when their drawing advantages were controlled for. Here artists’ advantage was only 0.14 SDs, a quite small effect size, $F(1,42) = 2.28, p > 0.10$. Consistent with (APA2), this result suggests that technical proficiency in drawing contributes to artists’ perceptual advantages.

Kozbelt’s findings are consistent with other research. Cohen and Bennett (1997) identified four potential sources for differences between artists’ and nonartists’ drawings: misperceiving the object to be drawn, poor motor coordination, making poor decisions about which features to include in the drawing (or how to render them), and misperceiving one’s emerging drawing. They found little evidence for the latter three possibilities and concluded that the primary reason nonartists cannot draw what they see is that they misperceive the object to be drawn. They argued as a result that artists’ drawing skill is grounded in an unspecified set of perceptual skills. Mitchell et al. (2005) found that participants’ misperception of the relative size of the two figures in Shepard illusions was more pronounced when they categorized the figures as tables rather than simple parallelograms (Figure 2). Further, there was a positive correlation between participants’ drawing errors and their reported degree of perceptual distortion. These investigations document the perceptual basis of nonartists’ limited drawing ability and imply that artists can overcome these limitations by superior perceptual skill.

The goal of this article is to motivate a model that explains Kozbelt’s results and supports (APA1) and (APA2). We hypothesize that, as artists develop technical proficiency, they develop a novel class of knowledge that defines artworks in their medium as a particular category of objects. We argue that this knowledge is encoded in two ways: (1) as spatial schemata that represent sets of stimulus features sufficient
for accurate depiction, and (2) as motor plans for rendering these stimulus features in a medium. Recent research in cognitive neuroscience demonstrates that spatial schemata and motor plans serve as the grounds for complementary attentional strategies, which enhance the perceptual encoding of stimulus features diagnostic for the identity of objects and inhibit the perception of potential distractors (Kanwisher & Wojciulik, 2001; Schubotz & von Cramon, 2003). Therefore, we argue that artists develop novel attentional strategies, derived from a unique form of practical knowledge, that productively bias their perception to stimulus features sufficient for adequate depiction in their medium:

APA3. Artists’ technical proficiency in a medium confers an advantage in visual analysis, which consists of the ability to focus attention on sets of stimulus features sufficient for adequate depiction.

We focus on drawing as a paradigmatic example of the effects of technical proficiency in art on perception. One might object that the goals of contemporary art are sufficiently different from the goals of traditional art (Belting, 1987) to render this
view of the role of drawing obsolete. The goal of this article is to generate a model for the role of drawing skill in visual analysis. If our analysis does not yield a theory relevant to the production of contemporary art, we are content to limit the scope of the model.

However, the influence of practical biases in perception is not limited to depictive contexts. Artists across many media, styles, and genres employ drawings in the process of developing the formal and compositional structure of their work (e.g., sketches for painting, sculpture, and installation; gesture drawings for performance and dance; storyboards for film; etc.). Thus, drawing skill is useful in a broad range of aesthetic contexts, which is reflected in the fact that it is the central tool for teaching visual analytic skills in art schools. For instance, the 2006–2007 course description for Drawing I at the Rhode Island School of Design reads:

Drawing skills are fundamental to the artist’s visual education. Through guided practice, beginning students sharpen their powers of observation as they learn to translate what they see into drawings. Using basic tools and materials, students develop their drawing skills by exploring...line, form, value, proportion, perspective and composition. Through a series of graduated exercises, students learn to draw...with confidence and prepare for further work in a variety of art media. (Rhode Island School of Design, 2006, p. 1)¹

Nonetheless, we expect that drawing skill is just one aspect of technical proficiency that influences visual analysis. We hypothesize that artists’ perceptual advantages are subserved by a diverse set of perceptual strategies derived from the skills necessary for artistic production in a broad range of media.

It is also important to emphasize that this is neither a theory of art, a theory of aesthetic response, nor a theory about how spectators perceive artworks. Rather it is a model for understanding the psychological processes underlying artistic production. However, we do expect that the model can contribute to explanations of art historical and aesthetic phenomena. For instance, the shift in focus in Western art away from traditional representational goals has been attributed to the explosion of styles in the Modern period (Belting, 1987). This explosion is a consequence of a shift in attention from optics and the structure of depicted space to medium-specific constraints on the formal and compositional structure of artworks.² Our model is derived from the observation that object recognition is a product of the conjunction of sensory inputs and category knowledge (see Palmer, 1999, p. 58; Winner, 1982, pp. 99–100). Various sets of formal marks are sufficient to enable a viewer to categorize, and so perceive, a stimulus as a particular scene or object. Thus, in principle, an unbounded set of different types of marks or styles are sufficient to induce the perception of a depicted scene or object in a drawing or painting. This, in turn, entails that the explosion of style in the Modern period is a natural consequence of the shift in attention from the structure of depicted space to the formal constraints on artistic practice in a medium. Therefore, our model generalizes to the explanation of some art historical phenomena.

Finally, the choice of formal vocabulary and style is an indicator of the semantic intentions of the artist (Rollins, 2004). The ability to categorize a work of art as
exemplifying a particular style is essential to both recovering this information from the surface of a painting and understanding the symbolic value of its representational and formal features. We therefore expect that the same attentional processes that constrain artists’ perception of stimuli in productive contexts also constrain spectators’ perception of artworks in aesthetic contexts. If this assumption is sound, then our model also helps to explain the roles of art historical knowledge and interpretation in aesthetic contexts.

In Section 2 below, we introduce two variants of the standard view of the influence of artists’ methods in visual analysis (Fry, 1909/1981; Gombrich, 1960). In Section 3, we relate this view of artists’ practice to contemporary research in cognitive neuroscience on the role of memory and attention in ordinary perception. In Section 4, we discuss evidence that motor plans play a complementary role to spatial schemata in visual analysis. We argue that this provides a mechanism to link proficiency with the tools of a medium to artists’ perceptual advantages. Finally in Section 5, we summarize the model and suggest some ways to test it.

2. Fry–Ruskin and Gombrich on Artists’ Methods

The view that artists’ formal studies and techniques are a means for overcoming the biasing effects of knowledge in perception is exemplified in the writing of Roger Fry and E.H. Gombrich (see also Cohen, 2005; Cohen & Bennett, 1997; Mitchell et al., 2005; Winner, 1982). Fry (1909/1981) asserted that the biological function of vision is to provide data for practical action (pp. 13–15). Familiarity and practical necessity cause the functionally salient attributes of an object’s appearance to become labeled by the visual system. Once this occurs, perceivers attend only to an object’s label and ignore its actual appearance in a context (e.g., we ignore the actual phenomenal, elliptical shape of a dinner plate on a table in front of us and perceive it instead as a flat circular object in depth). Therefore, Fry argued that one’s knowledge of the ordinary shapes and functions of types of objects interferes with perception such that the actual structure of an object’s appearance is rendered invisible to a viewer.

Picasso’s ‘Baboon and Young’ can be used to illustrate this phenomenon. Picasso cast this sculpture from found objects: the head is cast from a pair of toy cars, the body from a clay pot, and the tail from a truck spring. However, the outline of the cars roughly resembles the shape of a primate’s head, and the relative positions of the windshield, hood, grill, and bumper of the top car echo those of a primate’s eyes, nose, and mouth. When presented with a black-and-white image of this figure, students rarely notice that its head is cast from toy cars, even when it is projected 15 feet tall on an auditorium screen for extended periods of time. Several colleagues have reported the same experience in front of the sculpture itself at MoMA. This anecdotal evidence suggests that naïve subjects misperceive images of the sculpture because they attend to the rough shape of the figure (e.g., the outline of the ‘head’ and the general configuration of features constitutive of its ‘face’), and ignore the
image features critical to its identity as a car (e.g., the windshield (eyes), grill (nose), and bumper (mouth)). Therefore, familiarity with the identity of the sculpture’s subject, a baboon, renders the actual shape of the stimulus invisible to ordinary viewers.

Fry argued that artists’ methods (e.g., the development of drawing skill) ground perceptual strategies that enable them to ‘see-past’ the influence of cognition in vision to the veridical ‘structure of appearances’. Gombrich noted that this view rests on a logical error: if the role of practical knowledge in perception is the product of psychological processes constitutive of the biological function of vision, then cognition plays an intrinsic role in perception that cannot be overridden (Gombrich, 1960, pp. 297–299). Instead, artists’ formal methods are means to discover and harness the influence of practical knowledge in perception. Further, Gombrich argued that historical differences between artistic styles (e.g., Hudson River School vs. Photorealist landscape paintings) demonstrate that there is no fixed set of formal cues necessary for adequate depiction. This demonstrates that artists do not aim for literally accurate copies of stimuli. Rather they strive for systems of marks that suffice to induce desired perceptual effects. These formal vocabularies are developed by testing marks against artists’ own perceptual experience. Therefore, Gombrich conceptualized artists’ formal methods as strategies for analyzing the structure of perceptual experience, not the veridical structure of perceptual stimuli.

Consider the task of rendering a mountain landscape. Mathematical systems of perspective enable artists to place objects in the picture plane as they are located in the visual field. However, when mountains are rendered ‘accurately’ in two-dimensional media they can appear unrealistically small (e.g., vacation snapshots of panoramic vistas) (Gombrich, 1960, p. 311). The difficulty is that perceptual mechanisms that scale up the appearances of distant objects in the perception of three-dimensional scenes and objects do not adequately scale up the appearances of distant massive objects in two-dimensional images. This is not to say that pictures are not subject to the same perceptual transformations as occur when viewing a three-dimensional scene. We have argued that the formal features of paintings and drawings suffice to induce the perception of depicted scenes and objects because they trigger the visual processes responsible for ordinary vision. Rather, depth cues from binocular rivalry and motion parallax are unavailable in the two-dimensional surface of pictures. Artists thus learn to fudge the relative size and position of distant objects like mountains in order to more accurately approximate their apparent mass and scale. For instance, Bierstadt severely foreshortened the foregrounds and exaggerated the heights of the mountains relative to the width of their bases in his paintings. Geographers have formalized this depictive strategy into a system for drawing sections of mountain ranges (Gombrich, 1960, p. 311). Kubovy (1986, pp. 104–121) discusses analogous Renaissance strategies for handling marginal distortions and the depiction of spheres, columns, and figures in perspective. He argues that these strategies reflect ‘the perspectivist’s acceptance of the primacy of perception’ (Kubovy, 1986, p. 121).
Gombrich’s model suggests that artists’ methods represent a set of viewing strategies that are derived from the practical necessities of artistic production in a medium and define artworks as a novel category of perceptual stimuli. Thus, artists do not see-past the conceptual interference of practical knowledge. Rather, they ‘see-with’ a novel class of knowledge defining the structure and function of artworks in a medium. Gombrich argued that this type of artistic knowledge is encoded in schemata (e.g., abstract geometric patterns artists use to successfully depict the shapes and dynamics of objects and scenes in two-dimensional media). We interpret artists’ spatial schemata as structural descriptions (Kosslyn, 1996, pp. 216, 263; Palmer, 1975, p. 289; Solso, 2003, p. 224) that define the shapes of objects, and the placement of objects in scenes, relative to medium-specific constraints on artistic production.

Gombrich (1960, pp. 147–172) described ‘how-to’ manuals that provide instructions and simplified line drawings to train artists to the basic proportions of objects as evidence for artists’ use of schemata. For instance, in modeling a standing human figure in a dynamic pose it is critical to get the relative placement of the volume of the quadriceps and the calf muscle across the knee cap correct. The mass of the quadriceps is raised on the lower portion of the thigh in this context, and forms a roughly triangular shape whose apex points to the inside of the knee. The calf muscle is raised towards the knee and rolled to the outside of the shin. The relative downward and upward thrust of these two masses should form parallel vectors when viewed from the front (Lanteri, 1902/1985, p. 147). These types of schemata can influence perception in several ways. They direct attention to features and characteristics of an object or scene that might otherwise go unnoticed or unperceived. Further, they provide artists with frameworks for the relative placement of objects and their parts, which help locate their expected positions in the visual field (e.g., the central features of a face form an inverted equilateral triangle horizontally bisected by the midline of the head). In certain contexts the latter bias forms recognition by filtering out the effects of practical knowledge on the perception of an object (Cohen, 2005).

2.1. Summary

In this section we have presented art historical perspectives that exemplify the traditional view of the influence of artists’ methods in visual analysis. Fry and Gombrich disagreed about the epistemic object of artists’ formal techniques and methods. Fry contended that artists’ methods yield knowledge of the structure of the stimulus. Gombrich argued that they represent artists’ knowledge of the structure of phenomenal experience. However, both share a critical assumption: practical knowledge functions as a bias that constrains visual analysis in ordinary perception. These biases interfere with a perceiver’s ability to recover formal cues from the visual field sufficient to induce the perception of realistic three-dimensional scenes on a two-dimensional surface. This entails that artists must develop perceptual strategies to overcome their judgments about the identities and functionally salient features of
objects. Therefore, artists’ methods can be interpreted as sets of perceptual strategies that enhance performance in visual analysis.

3. A Diagnostic Framework for Perception

We propose a model for artists’ perceptual advantages (Figure 3) derived from Philippe Schyns’ (1998) diagnostic recognition framework for object identification. Theories of object identification have traditionally divided visual perception into two sets of processes: form recognition and object categorization. The standard computational argument for this division of labor rests on a logical point. Perceivers recognize the identity of objects in the visual field by matching visual representations of the geometry of a stimulus to stored records of the ordinary shapes and functions of object types. Therefore, form recognition must precede and operate independently from object categorization. However, researchers like Patrick Cavanaugh (1991), Schyns, and Stephen Kosslyn (1996) dispute this traditional computational model. They argue that object recognition does not require a full representation of the form

![Figure 3](image-url)

**Figure 3.** A Model for Premotor Contributions to Artists’ Perceptual Advantages.

*Notes:* Spatial schemata and motor plans play complementary roles in shifting selective attention. These mechanisms also enhance the encoding of expected diagnostic features in the visual field and inhibit the perception of local distracters: (a) sensory inputs are initially encoded in a visual buffer; (b) early visual processing culls diagnostic image features from sensory inputs encoded in the visual buffer; (c) object recognition involves matching diagnostic image features to category knowledge; (d) the closest match is used to generate a perceptual hypothesis about the expected identity of the stimulus in spatial working memory; (e) spatial schemata and information stored in long-term memory are used to prime the visual system and direct attention to expected diagnostic features; (f) motor plans serve as complimentary schemata to prime the visual system and direct attention to features diagnostic for anticipated actions.
of an object or the structure of a scene. Rather, the visual system can get by with incomplete sets of object cues that suffice to direct attention to diagnostic features in the visual field (see also Cohen, 2005, p. 999).

Diagnostic features are defined as sets of image cues that are sufficient to determine the identity of a stimulus. On this account, object identification is interpreted as a hypothesis testing process. Minimal sets of image features (e.g., coarse-grained visual patterns or sets of fine-grained features) are used to develop perceptual hypotheses about the most likely identity of a target object. These hypotheses, in turn, direct attention and prime the visual system to the expectation of confirming diagnostic features at particular locations in the visual field. Therefore, Cavanaugh (1991) and Kosslyn (1996) have argued that object identification is a bootstrapping process subserved by a distributed network of different resources.

3.1. Diagnosticity

Schyns’ identifies two critical variables in object recognition tasks: cue availability and cue diagnosticity. Cue availability is determined by both the perspective of a viewer and physiological and computational constraints on the availability of perceptual information to a particular system. Cue diagnosticity is determined by the information processing requirements of a particular task. Consider the two geometric patterns in Figure 4. Viewers familiar with a Necker Cube readily recognize that the two interior vertices in the left-hand figure, A and B, are diagnostic for two unique cubes at different orientations. Categorizing the figure as a Necker cube

![Figure 4](image)

**Figure 4.** Necker Cubes. Necker Cubes illustrate the claim that how one categorizes the content of an image can alter the diagnosticity of image cues, which can alter the way the input is organized into perceptual cues. Categorizing the left-hand image as a Necker Cube enables viewers to alter the relative spatial position of vertices A and B at will. Similarly, it enables viewers to perceive vertex C as either a single point in two dimensional space or as overlapping points in three-dimensional space in two different configurations. Therefore, how one categorizes a stimulus can influence how one perceives the stimulus.
enables viewers to alternately conceptualize vertex A or vertex B as the closest point in the foreground of the image. Viewers are subsequently able to both perceive the figure in depth and alter its perspectival projection by shifting attention from one junction to the other.

The right-hand pattern is harder to perceive as a Necker Cube. One must consciously concentrate on categorizing the figure as a cube in order to perceive vertex C as overlapping A and B junctions. Otherwise it defaults to a two-dimensional pattern of six equilateral triangles. Categorizing the pattern as a cube, therefore, alters the diagnosticity of C from the depiction of a single point in a two-dimensional image to the depiction of two overlapping points separated in depth. These examples demonstrate that how one categorizes an image can alter how one analyzes a stimulus into object cues, how one organizes these cues (e.g., altering whether vertex A or B is in the foreground), and even how one perceives them (e.g., as two corners in a three-dimensional space as opposed to one point in a two-dimensional space) (Schyns, 1998, p. 150).

Schyns’ diagnostic recognition framework can be used to explain the importance of specialized spatial schemata for artists’ perceptual and depictive practices. The salience of image features in a perceptual context depends on the particular perceptual hypothesis guiding attention. Artists’ spatial schemata encode knowledge of sets of image features sufficient for adequate depiction in their medium. In two-dimensional media, these sets of visual cues induce realistic three-dimensional perceptual experiences because they are image features that are diagnostic for the identity of the depicted scene or object. Therefore, schemata not only enable artists to recognize stimulus features sufficient to produce particular perceptual effects in their medium, but also confer perceptual advantages for visual analysis in ordinary perceptual contexts.

3.2. Kosslyn’s Hypothesis Testing Model

Kosslyn’s model for object recognition provides a neuropsychological explanation of the role of cue diagnosticity in perception. Kosslyn argues that object recognition is an example of cooperative computation (i.e., a set of interconnected subprocesses that share information and so function as a larger distributed system) (Kosslyn, 1996, p. 121). Cooperation is a property of opportunistic processing systems (Kosslyn & Koenig, 1995, pp. 76–77). The visual system is both cooperative and opportunistic in the sense that it exploits any and all resources available to it in order to promote fast and efficient recognition of the attributes necessary for cognition and action. Kosslyn has argued that this principle has two key correlates. First, coarse-grained visual patterns and minimal sets of fine-grained image features are ordinarily sufficient for basic object identification. Second, once an adequate match has been discovered between a set of image features and the stored record of an object’s shape or function, sensory information not needed for the continued operation of the system in that context is discarded. In this view, visual images are constructed piecemeal, beginning
with minimal sets of diagnostic features. Detail is ‘filled in’ only to the extent that it is needed for a given task.

For instance, coarse-grained features defining the general shapes of object types are ordinarily sufficient for object identification (e.g., the occlusion boundary and rough configuration of ‘facial features’ in black and white images of Picasso’s ‘Baboon and Young’). However, reaching and grasping may depend on a detailed representation of the structure of an object part. In the former case, there may be no need to attend to fine-grained details (e.g., the features that identify the baboon’s ‘eyes’ as the windshield of a toy car). In the latter case, information about the overall shape of the object may be superfluous to the task and can be discarded once the feature to be grasped has been located. In this regard, perceptual (object recognition), cognitive (object identification and task demands), and motor (reaching and grasping) systems cooperate to determine cue diagnosticity. The net result is an efficient perceptual system directed at the goals of the organism, which naturally filters out irrelevant sensory information. In this context, artists’ spatial schemata represent a novel form of practical knowledge that biases their perception to stimulus features sufficient for adequate artistic production in their chosen medium and style.

3.2.1. A four-stage model
There are four stages in Kosslyn’s hypothesis testing model for object identification. First, sensory inputs from each fixation are recorded in a short-term visual buffer. Second, a pattern recognition subsystem generates the basic formal structure of a visual image from sensory inputs in the buffer. Third, the output of the pattern recognition system is matched to prior knowledge of the general shapes and functions of objects in a categorization subsystem that involves associative memory. Fourth, if there is no good match between the output of the pattern recognition subsystem and stored records of the shapes of object types, the best match is used to generate a perceptual hypothesis about the form and identity of the scene or object represented in the current visual image. Perceptual hypotheses are, in turn, used by an attention shifting subsystem to generate expectations about the identity and location of potential diagnostic features and to shift attention accordingly (e.g., if a sculpture depicts a baboon, one expects to find a face with eyes and a mouth, not a windshield and bumper).

The critical stage of processing for our model is the attention shifting subsystem. Kosslyn argues that this subsystem exploits nonreciprocal connectivity between areas of prefrontal cortex associated with working memory and areas of the ventral stream associated with form recognition (Kosslyn, 1996, p. 251). One function of these connections is to instantiate low-level neural representations of objects or their parts in the inferior temporal and occipital cortex. This amplifies the sensitivity of the visual system to expected diagnostic features in the sensory inputs (Kosslyn, 1996, pp. 287–289). Therefore, the attention shifting subsystem tunes visual processing to the content of perceptual hypotheses.
3.2.2. Selective attention

Recent research on the role of memory and selective attention in visual perception supports Kosslyn’s model. There are feedback projections from areas of prefrontal cortex associated with working memory to areas TE and TEO of inferior temporal cortex associated with higher-level visual processing tasks like pattern and form recognition (Desimone & Duncan, 1995, p. 196), from TE and TEO to the areas V4 and MT in the occipital cortex responsible for culling color, form, and motion cues from retinal inputs (Kastner, 2004, pp. 147–148), and from the frontal and parietal eye fields (FEF/PEF), via the superior colliculus, to the lateral geniculate nucleus (LGN) in the thalamus (Awh & Jonides, 2001, pp. 119–226; Shipp, 2004, pp. 223–230). These feedback projections alter the baseline firing rates of target populations of neurons throughout the visual system prior to endogenous shifts of attention. Shifts in the baseline firing rates in these populations enhance the encoding of diagnostic image features, objects, and object parts at attended spatial locations.

This mechanism for attentional priming works in two ways. First, relative to a perceptual hypothesis, an excitatory response for a salient population of neurons increases the sensory signal associated with a target or location. In LGN and early visual cortex, this modulates the signal-to-noise ratio in the input stream (O’Connor, Fukui, Pinsk, & Kastner, 2002, p. 1206). Therefore, perceptual hypotheses bias visual analysis at the very earliest level of sensory processing. Second, a local inhibitory response specific to endogenous attention shifts filters out irrelevant features that are potential distractors. This effect is most pronounced in areas V4 and TEO (Kastner, 2004, p. 150). The activation of TEO cells whose receptive fields encompass two competing stimuli is ordinarily suppressed relative to the firing rates of the same cells for a single stimulus. However, single cell studies demonstrate that when primates covertly attend to one of these two stimuli, the activation of these cells matches their single stimulus firing rates. The net result is that attention serves as a filter to select sets of features salient for a particular perceptual hypothesis and to discard sensory information that is not. Therefore, the ascription of meaning to a visual image in object identification is a mechanism for directing selective attention and the way one attends to the visual field determines not only what one perceives, but also how one perceives it.

One consequence of this type of model is that visual recognition does not depend on a full representation of the global form of a scene or object. Rather, incomplete spatial information from each coherent glimpse of the visual field is shunted forward to associative memory and used to generate a perceptual hypothesis about the identity of the image. The only spatial information maintained online in working memory over iterations of the process is information salient to a particular hypothesis. This increases the efficiency of perceptual processing by ensuring that the visual system generates and maintains only the minimal set of image features necessary for recognition and action in a particular context (see also Cohen, 2005; Hayhoe, 2000). For instance, the text on someone’s tee shirt is not necessary for either the task of identifying them or reaching out to shake their hand. Sensory information defining the shapes of these letters can therefore, be discarded as spurious to those tasks without any loss of function.
3.3. Summary

In this section we have introduced a diagnostic framework for perception and presented evidence, which demonstrates that selective attention productively biases perception to sets of stimulus features sufficient for cognitive and practical tasks. Our model explains the failings of naive viewers’ visual analyses of Picasso’s ‘Baboon and Young’ as an artifact of ordinary perceptual processes subserving visual search and object recognition. The conjunction of the occlusion boundary of the figure and the coarse-grained visual pattern that defines its ‘face’ is sufficient to identify the sculpture as a baboon. Given the sculpture’s title, naive subjects simply fail to attend to the fine-grained image features that would override this perceptual hypothesis. Therefore, they literally perceive the figure as a baboon and fail to perceive it as a pair of toy cars balanced on a clay pot. Analogously, artists’ superior performance in visual analysis is the product of category knowledge, derived from media-specific technical proficiency and encoded as spatial schemata, which productively bias their perception to sets of stimulus features sufficient for object recognition.

Cohen (2005, pp. 1007–1008) suggests a similar mechanism for artists’ advantages in drawing tasks. He argues that novel attentional strategies enable artists to focus on stimulus features sufficient for accurate depiction. The deployment of these strategies generates virtual void viewing conditions that minimize context effects of stimulus interpretation in perception. Attentional strategies thereby enable artists to more accurately depict what they perceive. Our model provides a mechanism to explain how these attentional strategies would work. However, although the net effect is the same, our model differs slightly from Cohen’s. Following Gombrich, we interpret the effects of artists’ attentional strategies as due to context effects of novel category knowledge specific to drawing tasks. Therefore, we argue that artists’ productive strategies harness rather than inhibit the influence of stimulus interpretation in perception.

Finally, we have defined artists’ spatial schemata as structural descriptions that encode the general shapes of objects and scenes relative to medium-specific constraints on artistic production. Gombrich’s (1960) description of artistic ‘how-to’ manuals suggests that schemata are initially learned as declarative rules for the production of visual patterns. However, artists’ formal strategies are rarely restricted to the types declarative rules used to teach art. This suggests that artists also acquire spatial schemata through implicit processes as they discover new aspects of visual experience or productive techniques. Nonetheless, implicit perceptual learning requires selective attention to salient, diagnostic information. Therefore, learning is more efficient if implicit processes are bootstrapped by explicit knowledge, and declarative knowledge of basic proportions and drawing techniques should provide a perceptual advantage to trained artists in at least some contexts.

4. Motor Skill

We have argued that artists’ spatial schemata are perceptual strategies for recognizing and reproducing sets of diagnostic image cues. Kosslyn’s model of
object identification explains how this type of knowledge can influence perception and thus explains why artists are better at visual analysis and form recognition. In this section, we argue that Kosslyn’s model can be generalized to explain the influence of drawing skill in artists’ perceptual advantages. We hypothesize that, through extensive practice, the basic spatial information encoded in artists’ schemata becomes transferred to motor plans for the production of marks in a medium. Recent research demonstrates that motor plans play an analogous role to spatial schemata in selective visual attention. Therefore, Kosslyn’s attention shifting mechanism also explains how technical proficiency in a medium contributes to artists’ perceptual advantages. Eye movement research provides preliminary behavioral evidence in support of this claim. Tchalenko, Demere-Maroo, Hu, and Yang (2003) found that trained artists exhibit precise, repetitive paper-to-model-to-paper fixation patterns that are focused on to-be-drawn stimuli, uninterrupted by extraneous events, and closely coupled with hand movements. Further, they have demonstrated that this skill enhances voluntary eye control in nondrawing contexts (see also Cohen, 2005).

4.1. Motor Plans and Visual Attention

The premotor cortex (area 6) is associated with and specialized for motor intention based on prior motor plans. Area 6 is divided into two functional regions: the supplementary motor area (SMA), which is specialized for internally driven motor planning, and the premotor area (PM), which is specialized for externally stimulated motor preparation. Single cell studies of nonhuman primates and imaging studies of normal human participants demonstrate a functional division within both SMA and PM: a rostral region associated with planning and a caudal region associated with execution (Schubotz & von Cramon, 2003, p. 121). The rostral regions of SMA and PM are closely interconnected with area 46 in the dorsolateral prefrontal cortex and consists largely of neurons tuned for spatial orientation (Arikuni, Watanabe, & Kubota, 1988; Barbas & Pandya, 1987; Bates & Goldman-Rakic, 1993; Lebedev & Wise, 2001; Lu, Preston, & Strick, 1994). Area 46 is associated with spatial working memory and is hypothesized by Kosslyn (1996) to be critical for directing the influence of perceptual hypotheses on early visual processing. Further, the activation of these spatially tuned regions of SMA and PM is correlated with selective attention for locations, features, and objects (Awh & Jonides, 2001, p. 122; Schubotz & von Cramon, 2003, pp. 120–122). The rostral regions of SMA and PM are therefore, hypothesized to play a complimentary role to dIPFC in visual recognition. This yields the following functional hypothesis: motor plans are employed to direct attention to, and enhance the perception of, environmental features salient to an anticipated action (e.g., image features to be depicted while drawing).

Our model gains further support from research demonstrating that cells in rostral PM respond both to visual and tactile stimuli (Graziano, Yap, & Gross, 1994, p. 1054; Lebedev & Wise, 2001). These bimodal cells encode the locations and orientations of targets for reaching and grasping. Their visual receptive fields extend from the hand/arm location of their tactile field into the visual field and move relative to limb
orientation independent of gaze direction. Anticipatingly orienting one’s hand to a stimulus in preparation to draw an expected image feature at a location should, as a result, function as an independent means to amplify the baseline firing rates of the population of neurons in the visual cortex that encode that feature in the sensory input. Therefore, the anticipation of shifts in gaze direction and limb orientation in preparation for drawing a target should function as complementary means to enhance the perception of diagnostic features in the visual field.

4.2. Visuomotor Perceptual Strategies

The idiosyncratic perceptual strategies used by DF in copying tasks provide an example of the types of visuomotor attentional strategies that we argue underlie artists’ perceptual advantages. DF suffers from associative visual agnosia. Patients suffering from this type of agnosia have generally good cognitive abilities, can recognize objects via other perceptual modalities, and ordinarily have adequate low-level visual sensory capacities (Farah, 2004, p. 269; Milner & Goodale, 1995, pp. 125–126). However, they cannot recognize objects visually. Interestingly, these patients can often make accurate copies of line drawings even though they cannot visually recognize the subject of either the target or their copy. For instance, DF is quite good at copying the orientation of lines. However, she performs at chance in perceptual matching tasks with these stimuli indicating that she cannot visually discriminate their orientations (Dijkerman & Milner, 1997).

DF employed a curious strategy in preparation for copying tasks: drawing in the air in the line of sight between herself and the stimulus. When instructed not to do so, she would unconsciously draw in the air over the paper prior to marking her line. She reported that in both cases she was imagining tracing over the line while holding her pencil on the paper. She employed analogous strategies in manual orientation tasks (Milner & Goodale, 1995, pp. 140–144). DF’s description of these events suggests that she uses motor imagery as a strategy to compensate for perceptual categorization deficits responsible for her visual agnosia. Dijkerman and Milner tested this hypothesis by introducing tasks designed to interfere with motor imagery: counting backwards or tapping a finger from her nondominant hand while drawing. In both cases, DF’s copying performance fell to chance levels. When asked about her poor performance, she reported that the interference tasks occupied her mind too much for her to employ the imaginary tracing strategy. Dijkerman and Milner concluded that DF uses motor imagery in conjunction with hand movements to generate and test spatial representations of the orientation of the stimulus to be copied.

DF’s drawing strategy is notable for several reasons. First, Miall and Tchalenko (2001, p. 37) and Tchalenko et al. (2003, p. 718) reported that HO, a portrait artist they have studied extensively, employed similar tracing strategies that were closely coupled with directed fixations on to-be-drawn stimulus features. Second, Kosslyn (1996) and Jeannerod (1994) found that motor imagery recruits the same spatially tuned areas of SMA and PM involved in motor planning and directing selective attention for anticipated actions (see also Decety, 1996). Third, Kosslyn (1996)
presented evidence that the same mechanisms used to enhance the baseline firing rates of populations of neurons in selective attention are used to generate mental images. Finally, Schubotz and von Cramon (2003) reported that SMA and PM are involved in both spatial and object-directed attention. The conjunction of DF’s performance and this evidence suggests that she uses anticipatory motor planning to direct visual attention and enhance the perception of features in the visual field salient to copying the correct orientation of the line.

4.3. Motor Simulation and Apparent Distance

Our model supports Dijkerman and Milner’s interpretation of DF’s imaginary tracing strategy. Nonetheless, their experiment does not rule out the alternative hypothesis that DF’s performance is due to post-perceptual motor processing that facilitates copying (e.g., image-specific motor rehearsal or motor priming). One could test this possibility by evaluating whether imaginary tracing influences performance in perceptual discrimination tasks and, if so, whether motor interference eliminates these effects. Unfortunately, DF cannot consciously discriminate the orientations lines, so this method is unavailable in her case. However, Dennis Proffitt (2006) and colleagues have demonstrated that distance perception is affected by the task demands and goals of anticipated actions and that motor interference eliminates these effects. This research supports Dijkerman and Milner’s hypothesis and provides behavioral evidence for the influence of motor planning in perception.

Witt, Proffitt, and Epstein (2005) projected a target circle onto a table that was within reach of participants only if they held a baton. Participants either reached to the target with the baton or extended their arms and pointed to it with their dominant hand. They then adjusted the distance in the fronto-parallel plane between two comparison circles to match the apparent egocentric distance to the target. Finally, after the circles were removed, they indicated where the target had been by either reaching or pointing again. Results showed that the target appeared closer when participants anticipated reaching with the baton than when they anticipated reaching (pointing) with their fingers. Witt and Proffitt (in press) found similar effects when participants anticipated reaching (but did not actually reach) with the baton set beside them on the table and when they were asked to simply imagine reaching with the baton. There was no effect on apparent distance when participants held the baton but did not anticipate reaching.

This research informs our model in several ways. First, explicitly anticipating an action is a canonical example of motor planning. Proffitt and colleagues have demonstrated that anticipating actions scale the optical information in sensory inputs to a perceiver’s ability to perform that action (e.g., distances appear longer and hills appear steeper if one is wearing a heavy backpack) (Proffitt, 2006). This provides direct behavioral evidence that motor planning affects the spatial structure of perception. Second, Witt and Proffitt (in press) argue that the mechanism responsible for these perceptual effects is motor simulation, which involves covertly
or explicitly imagining performing an action. Motor imagery, as discussed above, recruits the same areas of SMA and PM that are involved in motor planning and directing selective attention.\textsuperscript{15} Thus, if motor simulation is responsible for action-specific influences in perception, then performing a concurrent task that engages these same mechanisms should eliminate perceptual effects. Witt and Proffitt (in press) tested this hypothesis by having participants engage in the same tasks as in Witt et al. (2005) while squeezing a rubber ball (cf Dijkerman & Milner, 1997). Doing so eliminated the difference between the reaching and pointing distance measures, providing further behavioral evidence that motor planning influences perception.

4.4. Summary

In this section we have provided neurophysiological and behavioral evidence to support our claim that motor plans play a complementary role to spatial schemata in selective attention that biases perception to stimulus features diagnostic for an anticipated task. We have also provided behavioral evidence suggesting the practical efficacy of these processes in drawing contexts.

5. Conclusions

Our diagnostic recognition model for artists’ perceptual advantages is derived from, and explains, the three claims detailed in Section 1. We argue that as artists learn the skills necessary for drawing, they learn to categorize visual stimuli as to-be-drawn scenes and objects. In doing so, they develop two types of strategies that bias perception to stimulus features sufficient for artistic production: novel spatial schemata representing sets of stimulus features sufficient for depiction and motor plans for rendering them in a medium. Recent research in cognitive neuroscience demonstrates that these strategies each independently enhance the encoding of diagnostic features perception. Therefore, conceptualizing a scene or object as a to-be-drawn stimulus, enhances visual analysis (APA1). Further, the complementary roles played by spatial schemata and motor plans in sensory processing explain Kozbelt’s (2001) findings that artists outperform nonartists on visual analysis tasks and that artists’ perceptual advantages are dependent on drawing skill. Therefore, the model provides a mechanism that explains the relationship between artists’ advantages in drawing and visual analysis (APA2). Finally, given the constructive role of selective attention in perception, this enables artists to perceive stimulus features that nonartists fail to perceive in ordinary contexts (APA3). The diagnostic framework we have outlined should not, however, be considered unique to artists. We predict that spatial schemata and motor plans germane to ordinary practice will be measurably productive factors in all perceptual contexts. Therefore, we expect that our model will generalize to explain other cases of the influence of expert knowledge in perception. For instance, batting experts must generate efficient visuomotor responses from specific perceptual cues (e.g., ball spin, arm angle, body posture).
Therefore, batting skill ought to confer an analogous advantage in visual analysis. Recent studies support this hypothesis (Land & McLeod, 2000).

Although there has been discussion of the role of spatial schemata in artistic production (Gombrich, 1960; Solso, 2003), our integration of motor skill and spatial schemata represents a novel contribution to the literature. We propose that our hypotheses can be tested by examining a population intermediate to artists and nonartists, such as art historians. Art historians study the history of techniques of artistic production and develop declarative knowledge of artists’ schemata and the means for deploying them, independent of accompanying motor plans or practice. We predict that art historians’ performance in drawing tasks should roughly resemble that of nonartists, demonstrating that they lack the motor skills necessary for adequate depiction. However, their performance in perception tasks should fall between artists and nonartists, demonstrating that spatial schemata contribute to, but do not alone account for, artists’ perceptual advantages.

Eye movement studies can also be used to test our model. Artists’ novel spatial schemata help focus their attention on the structure of a stimulus independent of its identity (see also Cohen, 2005). Artists’ viewing strategies should, as a result, help locate difficult to recognize diagnostic information in degraded images. We predict that artists and nonartists’ eye movements while looking at degraded stimuli should differ substantially. Further, since art historians share knowledge of artists’ spatial schemata, they should show scanning patterns similar to those of artists. Our prediction about artists and nonartists is consistent with results demonstrating that artists show more organized eye movements while drawing than when not drawing, and that nonartists’ eye movements are less systematic and do not differ between drawing and not drawing (Miall & Tchalenko, 2001; Tchalenko et al., 2003).

In sum, we argue that artists develop specialized spatial schemata and related motor plans that guide attention and enhance the perception of stimulus features diagnostic for the identities of objects and scenes in ordinary contexts. These two classes of specialized, expert knowledge ground perceptual strategies explain that artists’ advantages in visual analysis and form recognition. Further, evidence from the cognitive neuroscience of attention demonstrates that the deployment of selective attention both enhances the perception of diagnostic features and inhibits the perception of distracters as early as LGN. Therefore, the relative performance of artists and nonartists in visual analysis tasks indicate genuine perceptual differences. Naive viewers do not simply fail to notice the fine-grained image features constitutive of the windshield, grill, and bumper of Picasso’s baboon. They fail to perceive them at all. However, our model does not entail that sustained looking could not yield the same results as artists’ perceptual strategies. Nor does it entail that nonartists could not develop different strategies that would have the same net effect in visual analysis. Rather the directed perceptual strategies of artists are more efficient than either sustained looking or other, less focused, perceptual strategies. Therefore, although artists perceive the world differently than nonartists, this is not the product of an innate capacity. Rather, it is a matter of the selective deployment of learned
perceptual strategies originally derived from, and dedicated to, the processes of artistic production.

Notes

[1] See also Ruskin (1857), from the quote on p. 150.

[2] For instance, the qualities of the brush and viscosity of the medium constrain the types of fine-grained detail that can be accurately depicted in a painting. Degas’ use of mixed media (e.g., tempera and pastel crayon) for the dress in ‘Dancer with Bouquet, End of the Arabesque’ (1877) is a formal means to overcome these constraints (Rouart, 1988).

[3] An image of this work can be seen on the MoMA website (retrieved March 10, 2007: http://www.moma.org


[6] Akins (1996, pp. 337–372) makes a distinction between information carried in a signal and information encoded. Information is only encoded in a signal relative to the capacity of a system to read it. Information encoded in a signal is a subset of the information the signal carries (e.g., the light energy striking the retina carries infrared information that the human visual system cannot read).


[8] FEF and SC are primarily identified as mechanisms for directing eye movements. However, these areas are also involved in the attentional modulation of LGN.

[9] Endogenous attention shifts are internally, or cognitively, stimulated in contrast to exogenous attention shifts that are triggered externally, or by the stimulus itself.

[10] An additional source of support for our model comes from literature on inattentional blindness and changeblindness (see Chun & Marois, 2002; Triesch et al., 2003). Limited space precludes a discussion of this material and its relation to our model here.

[11] This functional distinction is sufficiently robust in SMA that the rostral area is often identified as a separate area, ‘pre-SMA’.


[13] ‘DF’ is the standard way of referring to Milner and Goodale’s ‘patient DF’. Milner and Goodale refer to her by these initials in all of their papers about her behavioral deficits and performance (see Dijkerman & Milner, 1997; Milner & Goodale, 1995). We are not familiar with her actual name.

[14] ‘HO’ is the standard way that Miall and Tchalenko refer to this artist in their papers. In fact, Tchalenko et al. (2003) refer to all the artists that they studied in their paper by their initials. The artist’s full name is Humphrey Ocean.


References


