What is a Drawing

Daniele Zavagno

Drawing, which fundamentally underpins all forms of graphic representation, is one of the oldest modes of expression, with the earliest recorded marks dating back approximately 30,000 years [Massironi 2002]. As a means of expression and communication, it predates and serves as a precursor to written language. It is also one of the most versatile and virtually universal expressive media, capable of transcending linguistic barriers. However, between the 1930s and 1960s, a variety of scholars argued that the recognition and understanding of what is represented in two-dimensional images –including drawings– is a learned skill. While this may hold true for scientific or technical notations, it is doubtful whether it applies to all forms of graphic representation. Yet, the thesis that we "learn to see" drawings is supported by several cross-cultural studies conducted with populations lacking a figurative culture. Participants were asked to recognize scenes and objects in photographs and line drawings, or

This article was written upon invitation to frame the topic, not submitted to anonymous review, published under the editorial director's responsibility.

to produce drawings themselves [e.g., Hudson 1967; Mundy-Castle 1966; Nissen, Machover, Kinder 1935; Segall, Campbell, Herskovits 1966; for a critical review, see Kennedy 1974]. A popular argument was that the meaning of spatial relationships in pictorial images is a learned construct [Stone, Church 1968]. Other research, however, supports the exact opposite. One of the most compelling studies was conducted by Julian Hochberg and Virginia Brooks [1962] using their own son as a subject. From birth, they limited his exposure to two-dimensional images. At 19 months, they began a test showing him line drawings of objects semantically familiar to the child, followed by photographs of the same objects. The unequivocal result was that the child was able to name almost all the images shown, whether they were drawings or photographs. The ability to recognize what is represented in a two-dimensional image –for which James Jerome Gibson [1954; 1979] coined the term "picture perception"– thus appears to belong to the human cognitive toolkit. But could it be otherwise, considering that our ancestors were already depicting figures on cave walls?

Drawing is often perceived as the product of a specific talent or skill, whether artistic or technical. Yet almost everyone has drawn at some point in their life, regardless of proficiency. What, then, is a drawing and how does it work? The question may seem trivial, but the answer is rather complex and has several ramifications. I will attempt an answer in relation to the perception of three-dimensionality –that is, space.

Let us begin with the most superficial layer: a drawing is a human artifact, intentionally created for diverse purposes, including ludic, aesthetic, communicative, exploratory, and design-related ends, as well as combinations thereof. The aesthetic dimension is almost always present, given that the human being is an "aesthetic animal". Whatever its purpose, a drawing is always a representation of something, even when abstract. One might think that only scribbles and doodles are exceptions, as they often lack a specific representational intent, much like the early drawings of young children. However, children's scribbles are not ends in themselves; they are exercises in motor control –a ludic activity enjoyed through the marks left on the pictorial support. More often than not, children provide an interpretation or meaning only after the work is completed [Arnheim 1974]. One might say: *post hoc, ergo propter hoc*.

Every drawing defines, and therefore represents, a space. This applies even to absent-minded doodles, as any mark traced on a support –which we may define as the "pictorial surface"– alters its spatial qualities. These marks are spatial notations that can perceptually nullify the material flatness of the support. Indeed, it is extremely difficult to create a pictorial image that completely negates any impression of three-dimensionality and spatial depth; every mark on a potential pictorial surface can induce a sense of spatiality in the observer. What is even more fascinating is that any discrepancy on a physical surface could trigger experiences of spatiality that transcend the flatness of the surface itself –experiences that are pictorial in nature and typically defined as "resembling images" [Alberti 1999] or "*pareidolia*" [for a discussion on these terms, see Zavagno, Della Vedova 2025, note 12, p. 211]. Leonardo da Vinci was fully aware of this phenomenon: «I shall not omit to include among these precepts a new device for study, which, although it may seem trivial and almost laughable, is nevertheless of great utility in awakening the mind to various inventions. And this is: if you look at certain walls soiled with various stains, or at stones of different mixed colors, if you have to invent some setting, you may see therein similarities to various landscapes, adorned with mountains, rivers, rocks, trees, vast plains, valleys, and hills in diverse ways; you may also see battles, and rapid actions of figures, strange expressions of faces, and costumes, and an infinite number of things which you can reduce to good and integrated form» [Leonardo Da Vinci 2019, pp. 89,90; trans. by the author].

One of the most established theories explaining the ability to see things where they are not physically present is the cognitivist approach. According to this view, we learn to recognize the spatial relationships between ourselves, objects, and physical space. The organization of a visual scene is thought to be based on "cues": units of visual information that must be learned through experience. In the context of picture perception, these are called "pictorial cues", as they are information that can also be found in two-dimensional images.

The cognitivist hypothesis suggests that artists identified and learned how to reproduce critical visual information contained in the retinal image (the cues) correlated with real three-dimensional stimuli. The theory posits that, starting from the set of cues reproduced by an artist, unconscious inferential processes are activated in the observer on a probabilistic basis. These processes, governed by experience and species-specific genetic knowledge, aim to recreate the visual scene that best fits the "reality of the facts", what is most congruent with the distal stimulus (the physical reality from which the proximal stimulus is originated).
While the cognitivist approach to visual perception is the most widely represented in general psychology textbooks, it faces evident problems and fallacies. Let us examine just one macro-issue regarding drawings and bi-dimensional images in general: the theory maintains that recognizing what is represented depends on unconscious inferential processes where visual input (cues) from retinal stimulation is combined with the observer's prior cognitive information to produce an output (visual experience) congruent with the distal stimulus. But herein lies the problem: when looking at a drawing, the visual experience is never congruent with the distal stimulus. This is not only because the distal stimulus is materially flat (consisting of the support and the marks), but also because a drawing is inherently subject to the artist's style. No human-made image is an objective rendering of physical reality for two reasons: 1. style conflicts with the concept of verisimilitude; even in hyper-realism, a drawing expresses the creator's viewpoint; 2. we do not perceive physical reality "as it is", because knowledge of the external world is mediated by our sensory systems. In visual perception, we simultaneously see less and more than what physical reality offers: less, because we are not sensitive to all electromagnetic waves; more, because we perceive "meaning" and "relationships" even when they are physically fictitious, e.g., the launching effect studied by Albert Michotte [1946]. Wolfgang Metzger [1971] rightly defined physical reality as "metempirical", as it lies beyond our capacity for direct knowledge.
What is extraordinary –and often overlooked– is that even a simple line drawing can trigger an

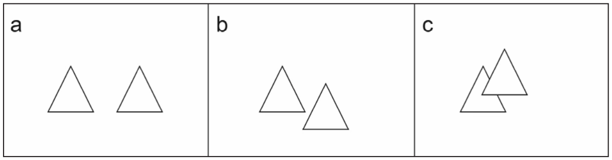
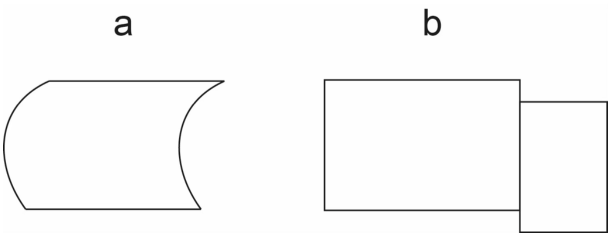


Fig. 1. a. most observers see the two triangles as coplanar; b. most observers see the two triangles as displaced at different distances; the perception of coplanarity is instead difficult to obtain in a stable way; c. example of partial occlusion that resolves the perceptual ambiguity of panel b concerning the spatial dislocation of the two triangles.

experience of three-dimensionality. In other words, the eye is remarkably easy to deceive. One could argue that we "see things" because pictorial cues have been placed on a surface either accidentally or intentionally. However, such an analysis risks missing a fundamental point: the ease with which the eye is "deceived" might signify that space is not a datum to be "reconstructed" from proximal (retinal) cues, but rather an intrinsic dimension of the visual system itself, regardless of the correspondence between proximal and distal stimuli. If anything, the role of these cues –which I prefer to call "indices" as a more theoretically neutral term– is to "shape" that space [Zavagno 2025] which, in the absence of proximal discrepancies, would result in the perception of an "ocular fog", i.e. "*Ganzfeld*" [Metzger 1930].
Every visual experience necessarily implies some experience of spatial three-dimensionality, even with materially flat images. Figure 1a shows two physically identical triangles aligned horizontally; there are four possible perceptual outcomes. The primary one is the perception of coplanarity; however, with minimal effort, one can alternately see the right or left triangle as being slightly deeper in space. Finally, one might perceive the triangles as holes. Figure 1b shows identical triangles but displaced in a different way on the pictorial plane: the impression of coplanarity is now virtually impossible, and the observer tends to see one triangle in the foreground and the other in the distance.
Creating a perceptually flat drawing –one that yields no impression of depth– is nearly impossible. Even in the coplanar case (fig. 1a), the triangles are seen

Fig. 2. a. the configuration is seen either as convex or concave; in other words, it does not appear like a flat surface; b. the two partially juxtaposed rectangles can also appear as parts of a same folded rectangular structure, so that one part lies on the sheet while the other is raised.



as lying against a homogeneous white background. "Partial occlusion" is a crucial index always present in the visual field (except in a *Ganzfeld*) that highlights the hierarchy of spatial planes (fig. 1c). One might suggest, then, that the challenge is not nullifying the flatness of the surface, but generating figures that appear three-dimensionally solid.

The remarkable thing about line drawings is that, despite their visual "poverty", they are highly informative and capable of inducing a perception of "solid" three-dimensionality. Figure 2 shows two multistable configurations. Figure 2a is usually seen as a sagittal section of a tube, showing its external convexity. However, with little effort, it can be seen as concave. Figure 2b is even more ambiguous, showing two rectangles with an adjacent side but slightly offset in height. Many readers may initially see coplanarity; however, the longer one observes it, the more likely it is to appear like a folded card, sometimes with the smaller rectangle projecting forward, or vice versa.

Looking at figure 2, some might hastily conclude that a line drawing remains inherently ambiguous and can never truly "deceive" the observer. Regarding this last point, figure 3 reproduces a drawing used by John Kennedy [1974] in an experiment where participants were asked to draw something in the empty space between the children. Several people, while focusing on that central space, absent-mindedly attempted to pick up the drawn pencil. This demonstrates that, in a given context, even a line drawing can take on the characteristics of a *trompe l'œil*.

Ambiguity depends on a drawing's complexity and its chosen geometry. Figure 4 offers graphic solutions to resolve the ambiguities in figure 2. The geometry of space is fundamental. It is often taught that linear perspective faithfully simulates what happens on the retina (the proximal stimulus). This is true, but we can perceive solid bodies using other projective systems, such as axonometric projections. These are commonly used in technical manuals (e.g., furniture assembly instructions) because they maintain constant dimensions regardless of where parts are placement in pictorial space. In contrast, in linear perspective, two physically identical objects are drawn in different sizes depending on their pictorial depth. Which method is more "natural" or better suited to represent the world as we see it? In my view, both are consistent with our perceptual experience. Linear perspective is the translation of the optical laws governing how we receive optical information: it is particularly important for

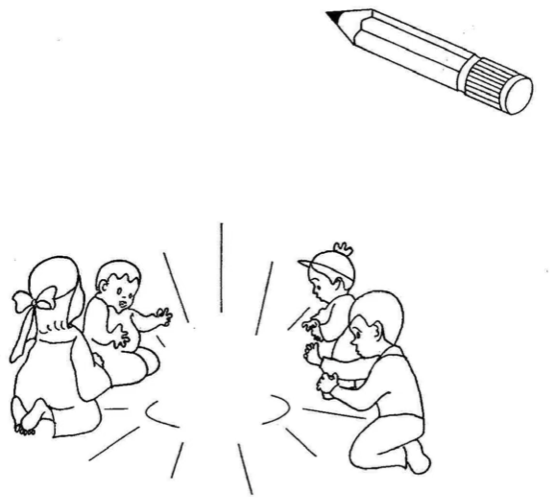
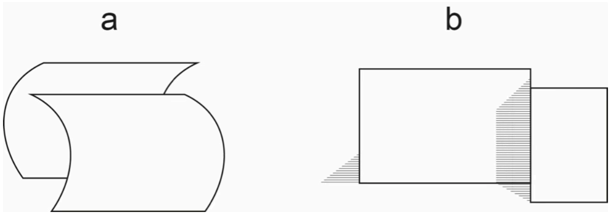


Fig. 3. Example of a figure used by Kennedy in an experiment in which the task was to draw the missing object in the center of the drawing. While focusing on the empty spot, several participants absentmindedly tried to pick up the pencil drawn above: Kennedy 1974, fig. 7, p. 52.

Fig. 4. Possible solutions to the multistable perceptual outcomes in fig. 2. The shading shown is not based on a geometrically constructed method; rather, it is intended solely to illustrate how a general use of shading can induce the perception of depth and three-dimensionality.



size constancy, which is provided by relationships in terms of perceived distances. In fact, according to Emmert's law [1881], the perceived size of an object depends on the product of two factors: the size of its retinal image (angular size) multiplied by the perceived distance. It is by virtue of this relationship that we do not perceive objects as enlarging or shrinking in proportion to their retinal projection, which, in fact, varies in proportion to the distance between us and the observed object. Axonometric projections, on the other hand, while not adhering to the laws of optics, provide us with perceptual experiences similar to the so-called visual concepts discussed by Arnheim [1974], whereby boxes appear as more or less regular parallelepipeds and soccer balls as circles, regardless of the angle from which we view them.

Both systems are valid as they both exploit the underlying architecture of our perceptual system. Could it be otherwise, given that humanity, in the attempt to translate perceptual experiences into images, has worked to refine these methods of representation over millennia?

Author

Daniele Zavagno, Department of Psychology, Università degli Studi di Milano-Bicocca, daniele.zavagno@unimib.it

Reference List

Alberti, L.B. (1999). *De Statua*. M. Spinetti (a cura di). Napoli: Liguori.

Arnheim, R. (1974). *Art and visual perception. A Psychology of the creative eye*. Berkeley: University of California Press.

Emmert, E. (1881). Größenverhältnisse der Nachbilder. In *Klinische Monatsblätter für Augenheilkunde und für augenärztliche Fortbildung*, n. 19, pp. 443-450.

Leonardo da Vinci. (2019). *Libro di Pittura*. M.T. Fiorio (a cura di). Milano: Abscondita.

Gibson, J.J. (1954). A theory of pictorial perception. In *Audio-Visual Communications Review*, n. 1, pp. 3-23.

Gibson, J.J. (1979). *The ecological approach to visual perception*. Hillsdale: Lawrence Erlbaum Associates.

Hochberg, J., Brooks, V. (1962). Pictorial recognition as an unlearned ability: A study of one child's performance. In *The American Journal of Psychology*, n. 74, pp. 624-628.

Hudson, W. (1967). The Study of the Problem of Pictorial Perception Among Unacculturated Groups. In *International Journal of Psychology*, n. 2, pp. 89-107.

Kennedy, J.M. (1974). *A psychology of picture perception*. San Francisco: Jossey-Bass Publishers.

Massironi, M. (2002). *The psychology of graphic images. Seeing, drawing, communicating*. Mahwah: Lawrence Erlbaum Associates.

Metzger, W. (1971). *I fondamenti della psicologia della gestalt*. Firenze: Giunti-Barbera.

Metzger, W. (1930). Optische Untersuchungen am Ganzfeld. II. Zur Phänomenologie des homogenen Ganzfelds. In *Psychologische Forschung*, XIII, pp. 6-29.

Michotte, A. (1946). *La perception de la causalité*. Louvain: Publications Universitaires de Louvain.

Mundy-Castle, A.C. (1966). Pictorial Depth Perception in Ghanaian

Children. In *International Journal of Psychology*, n. 1, pp. 289-300.

Nissen, H.W., Machover, S., Kinder, E.F. (1935). A study of performance tests given to a group of native negro children. In *British Journal of Psychology*, n. 25, pp. 308-355.

Segall, M.H., Campbell, D.T., Herskovits M.J. (1966). *The influence of culture on visual perception*. New York: Bobbs-Merrill.

Stone, L.J., Church, J. (1968). *Childhood and adolescence*. New York: Random House.

Zavagno, D., Della Vedova, D. (2025). *Fenomenologia della luce. Tra percezione e arte*. Milano: Mondadori Università.