Abstract - Affective - Multimodal

Interaction between Medium and Perception of Moving Images from the Viewpoint of Cassirer’s, Langer’s and Krois’ Embodiment Theories

Martina Sauer

Abstract

Everyday media consumption leaves no doubt that the perception of moving images from various media is characterized by experience and understanding. Corresponding research in this field has shown that the stimulus patterns flooding in on us are not only processed mentally, but also bodily. Building on this, the following study argues that incoming stimuli are processed not only visually, but multimodally, with all senses, and moreover affectively. The classical binding of a sensory organ to a medium, on whose delimitation the common understanding of multimodality is based, is thus abolished. Can arguments for this assumption be found? Building on the research of cultural anthropologist Ernst Cassirer and his successors, the two American philosophers Susanne K. Langer and John M. Krois, and supported by research from the life sciences in which, in parallel, the arts also play an important role in understanding image formation, as in the work of Heinz Werner and, following on from this, in the more recent research of the child psychologist Daniel N. Stern in collaboration with mirror neuron research by the group of neuroscientists around Giacomo Rizzolatti in Parma, justifications for this will be provided. What the multimodal processing of film images is concretely based on will finally be made clear by means of an exemplary analysis of a video film by the Swiss artist Pipilotti Rist. It aims to prove, that the prerequisites for the thesis can only lie in the fact, that both the modes of perception (visual, tactile, auditory, olfactory, gustatory) as well as the medially conditioned modes of image formation are based on non-discursive, formal-abstract structures that are also processed affectively. At the same time, the analysis thus shows how the selection and composition of the vital experienced forms can provide information about the socio-culturally shaped self-image of the producer and the viewer.
Derivation

The virulent discussion about the question of the connection between perception and image formation at the beginning of the 20th century can be observed in various research perspectives that dealt with it without mentioning the concept of medium or multimodality. Rather, the question of how the world is perceived was examined by means of (mostly painted) pictures. In this regard, the initial focus was on the mode of image formation, which was considered to be informative for this purpose. The fact that it not only addresses the visual, but above all the tactile perceptual ability was a central starting point. The discussion was further stimulated by the fact that, in addition to the obvious multimodality, the affective capacity of the recipient was also addressed. It quickly became clear that in addition to the mental processing of the composition, which aims to give it meaning, there is also a physical processing. It was precisely the aesthetically, bodily sensory arousal potential of the works that stimulated reflection. How can we understand the experience, the felt liveliness that is able to flow through the recipient of a work? Do the works have to be configured in a certain way in order to appear meaningful and, moreover, to appeal to different senses as well as to arouse sensations? But if this is so, what is the relationship between the experience of images and the perception of the world? Can a difference be shown between the two? Are the mechanisms of perception not universal? Alongside the sciences of art and philosophy, the life sciences have thus made an important contribution to the clarification of these questions from the very beginning, be it psychological, biological and zoological or neuroscientific research results. Already these early researches opened up that both, the perception of an image and the perception of the world, have to operate according to the same rules. It was implicitly assumed that there must be an analogy between the perception of worldly and creative aspects. What does it consist of? What connects worldly and medially mediated phenomena in such a way that they can be perceived, felt and understood with all senses?

Remarkably, it is precisely the early research on these questions, which remain open to this day, that I consider to be able to contribute to their clarification. Their focus on organizational forms can be seen crucial in this respect. They already emphasize that the principles according to which they are organized are purely formal and to that extent abstract. This basic idea will be adhered to in what follows. For from it can be derived according to the thesis put forward here, the common ground that can be regarded as the precondition of the analogy between worldly and medially conditioned perception.
Against this backdrop, it is assumed that the non-discursive mode of appearance of formal-abstract elements have no intrinsic meanings, as already Hans Jürgen Wulff states (Lexikon der Filmbegriffe: diskursiv/präsentativ, lastest update 2022). The proof that basically every medium is based on such non-discursive and thus abstract-formal elements has been elaborated by the philosopher Susanne K. Langer in four successive writings ([1953] 1967, 103, 369, 372). In this she coincides with research in developmental psychology, presented since the beginning of the century by the influential researcher Heinz Werner in 1926 and 1940, also confirmed by the current research of the American Daniel N. Stern since the 1980s, according to which all senses and, beyond that, the affective capacity of the recipient can potentially be addressed through them. In this sense, formal-abstract structures are not only multimodal, but also multimodal or basically amodal in nature (Werner [1926] 1959, 66-7, [1940] 1957, 96–7; Stern 1985, 47–68, 2010, [1985] 1992, 74–103, 2011). Against this background, it is natural to assume, as first advanced by the cultural anthropologist Ernst Cassirer, among others, in his exchange with Werner, and as taken up by the two pragmatist-influenced researchers Susanne K. Langer and John M. Krois, that the perception of world is also guided by the same principles (Sauer 2014b).

Research Background

In retrospect, it seems like a unique historical event how the University of Hamburg, newly founded in 1919, created a space in which researchers from different disciplines pursued the most pressing questions of the time together in close exchange, if only for a short time, until most of them had to emigrate with Hitler’s seizure of power in 1933. One central question that researchers have asked themselves can certainly be seen as the one that arose in the wake of Darwinian research: what distinguishes humans from animals? (Hartung [2001] 2004, 11–82). In addition to the human being’s power to form images and to act, it is above all in-depth questions of perception theory that have become the focus of common interest with regard to the formation of consciousness. Through them, the interests of the philosopher Ernst Cassirer met with those of the art and cultural scientist Aby M. Warburg and of the art historian Erwin Panofsky, but also with the research of the biologist Jacob von Uexküll, the psychologist Heinz Werner and the Bauhaus master Gertrud Grunow, to name just a few.2
Remarkably, their questions are primarily related to those from art history or formal aesthetics, in which the relationship between modes of perception and modes of image formation has been focus of discussion (Wiesing [1997] 2008, 25–205). Until the 1980s, their research was based on the concept that the particular view of the producer determined the method of image formation (Sauer 2016). Thus, an optical view of the world produces different images from a haptic one (Riegl 1901), a painterly one from a linear one, as can be seen in the example of Baroque and Renaissance artifacts (Wölflin [1915] 1923) and a dynamic (painterly) one from a representational (linear) one (Imdahl 1974, 325, 1987, 14–34, esp. 33). Only Gottfried Boehm, who founded the research center Eikones in Basel in 2006, departed from this specification, although he too explicitly followed Imdahl’s terminology and distinguished between processes of perception that take place simultaneously (visual, dynamic seeing) and successively (object seeing), to which, however, he did not assign any specific modes of image formation, be they painterly or linear (Boehm 1980, 120–22). This in turn has far-reaching consequences, since with the separation of perception and image formation no functional connection can be established on which a semiological and thus historically-culturally relevant theory could be built. This is also reflected in the research that follows Boehm’s theory of images (Alloa 2014, cf. review of Sauer 2015a) and in the far-reaching critique that his approach, but also that of Horst Bredekamp, has triggered, as will be shown below (Wiesing 2013, cf. review of Sauer 2014a).

In contrast to the approach of formal aesthetics, which was well known in the Hamburg circle (Panofsky [1932–1964] 1984, 187), the Hamburg researchers took a different approach. Basically, it was Ernst Cassirer who made the new considerations public with his main work, the three volumes of the Philosophy of Symbolic Forms published in 1923, 1924–25 and 1929. They are expressed in a theory of perception he developed, which is based on the premise that humans fundamentally do not perceive objectively, but are subjectively shaped by the affective reading of forms of motion and spatial forms; it forms the core of his theory of culture (Sauer 2008). Cassirer called this original perceptual activity expression-perception (Cassirer [1929] 1964, 94, [1929] 1979, 80). It concerns perception in general as well as image perception in particular, as he partly recorded in his late writing An Essay on Man, which he set down during the time of his American exile in 1942 ([1944] 1953, 190, [1944] 2007, 229). In distinction to formal-aesthetic assumptions, it can be stated, without this being explicitly addressed by Cassirer, that it is not intentional decisions to process the world be
it optically-dynamically or haptically-objectively, but that the differences in the stylistic conception of motifs (painterly and linear) observable over epochs are based on the premise that humans always and insofar directly experience the world as dynamically organized. This gives rise to the thesis—not to be developed further in the context of this study—that the variation in the stylistic elaboration then depends on the respective unconscious, culturally shaped mood.

For the question pursued here, which is interested in the interaction between perception and medium, it is rather essential to point out that Cassirer—just like formal aesthetics—assumes a functional connection between modes of perception and modes of images forming. According to both approaches, it is based on formal-abstract principles. But the one, what constitutes the perception in the world as dynamically-moving, are, according to Cassirer, not the objects as such, but precisely the forms of motion and spatial forms and thus also non-discursive, formal-abstract aspects. In the perception of the image, it is again abstract-formal elements and explicitly not the recognizable motifs that are grasped as “living forms” ([1944] 1953, 194, [1944] 2007, 233–34). Anthropologically conditioned, man is, as Cassirer put it following research from the life sciences in particular, determined by a “libidinal underlayer”, so that man always interprets everything as a dynamic event (Cassirer [1929] 1964, 78, 94, [1929] 1979, 66–7, 80). Anything that is perceived through the original form of perception, expressive perception, can therefore always be seen as symbolically significant (Cassirer [1929] 1964, 222–37, esp. 234–35, [1929] 1979, 191ff.). Aby M. Warburg, in particular, followed Cassirer in this view, creating, for example, a Mnemosyne Atlas, which collected the migrations of sensations, the so-called pathos formula, captured by expressive perception, as they are capable of expressing themselves, according to the initial thesis, in images of all cultures, and to this end compiled a comprehensive library (Böhme 1997). On the other hand, the methodological distinctions of the art historian Erwin Panofsky between phenomenon sense, meaning sense and document sense (sense of essence) or pre-iconographic, iconographic and iconological method ([1932–1964] 1984, [1939–1955] 1984) do not allow a direct connection to Cassirer’s theory of perception, even if he apparently oriented himself with it on the threefold division of the modes of consciousness that Cassirer introduced. According to them, the mythical, descriptive-aesthetic, and theoretical modes of consciousness allow man to distance himself from the original living mode of perception of the world (expression perception) through processes of detachment, without, however, ever losing the connection to it (Cassirer [1929]
1964, 103, [1929] 1979, 87). Thus Panofsky, too, initially emphasized the primarily perceptible layer of sense, the “phenomenal sense”, through which man grasps both the expression of what he sees and the factual sense. In contrast to Cassirer, however, he immediately inferred discrete emotions such as sad or happy and thus concrete impressions that we can get from a face, for example (Sauer 2020).

Cassirer, on the other hand, understands what is grasped by expressive perception as something much more indeterminate, namely only impulses and directions (forms of motion and spatial forms) that can be perceived as vigorous or weak (Cassirer [1929] 1964, 94, [1929] 1979, 80). Thus, Panofsky already loses touch with Cassirer’s distinctions with the introduction of the so-called pre-iconographic method. With the iconographic method, Panofsky then provided a set of instruments by means with which the stylistic differences and their corresponding meanings in the representation of a motif were comparatively recorded. In this way it was possible to create a catalogue of typologies, which is still used today, and in which, for example, the characteristic differences of the depictions Mary are recorded (Lexikon der christlichen Ikonographie, 1968–76).

In this methodological approach, analogies to the descriptive-aesthetic mode of consciousness become apparent. Points of reference to theoretical consciousness are furthermore evident in Panofsky’s iconological approach, through which the respective historical significance of the works can be developed by exploring their context. From a methodological point of view, Panofsky’s criteria of distinction have a pragmatic function in that, comparable to Cassirer’s approach, they can reveal the cultural-historical foundations of man, which he himself creates by means of symbolic forms (Cassirer [1929] 1964, 222–37, [1929] 1979, 191ff.; Panofsky [1939–55] 1984, 212). But just as formal aesthetics eventually loses its reference to the historical-cultural lifeworld, Panofsky’s iconological method, which is still valid for (cultural) visual studies and art history today, also tends to neglect the specificity of the image formation level. But just as formal aesthetics ultimately loses touch with the historical-cultural lifeworld (Prange 2004, 174–215, esp. 214–15), Panofsky’s iconological method, which is still valid for visual studies and art history today, tends to neglect the specificity of the formal level (Imdahl 1979, 14–5; Boehm [1978] 1985, 452–53). This, in turn, fundamentally contradicts the approach of Cassirer, who, however, only in his late writing, i.e. after the appearance of Panofsky’s first essays on the subject, explicitly stated that the perception of the recipient is ignited precisely by the formal-abstract elements, the “living forms” of the artworks, and in this way reality conveys reality via the works in a heightened
manner (Cassirer [1944] 1953, 194, [1944] 2007, 233). Indirectly, however, this approach has already been presented in the third, summary volume on the philosophy of symbolic forms:

Within this horizon, actually, the expressive perception is not only psychologically earlier than the perception of things .... It has its specific form, its own essentiality, which cannot be described, much replaced, by categories valid for other regions of being and meaning. ... The linguistic designation of movement, for example, almost always discloses this factor: instead of describing form of the movement as such, as the form of an objective spatial temporal process, language names and fixates the condition of which movement in question is an expression. “Quick,” “slow,” and if even “abrupt”—writes Klages, ...—may be understood in terms of pure mathematics; but “hurried,” “restrained,” “circumspect,” “exaggerated” are just names for conditions of life as for kinds of movement, and describe the latter by indicating their characters. Anyone who to characterize forms of motion and spatial forms, finds himself pectedly entangled in a characterization of psychic attributes, forms and movements have been experienced as psychic phenomena before they are judged by the understanding from the standpoint objectivity, and because language can express objective concepts through the mediation of the experience of impressions. (Cassirer [1929] 1979, 80, [1929] 1964, 94)

Accordingly, Cassirer characterizes this original form of perceiving as one that is characterized by experiencing and suffering. What is apprehended thereby receives an expression (Cassirer [1929] 1964, 88, [1929] 1979, 75). What is grasped in this way thus acquires a direction of meaning.

Against the backdrop of the ambivalent reception of Cassirer’s approach in the art sciences, it is not these who take up the theory of perception in the immediate aftermath. This can also be blamed on Cassirer himself, since he no longer explicitly addresses the approach that is fundamental to his philosophy in his late writings. Thus, due to Cassirer’s emigration via several stations to the US, it is the two American researchers who stand in the tradition of pragmatism, who follow up to him. Among them, starting in the 1940s, are the Whitehead’s student Susanne K. Langer and, much later, the co-founder of Bildakt at the
Humboldt University in Berlin John Michael Krois, also from America, who refers fundamentally on Peirce (Sauer 2014b). Even before Cassirer’s late work appeared in America in 1944, Langer building on the latter’s theory of perception and that of Whitehead, developed her own approach to art theory, which she published in 1942 and systematically expanded in further writings in 1953, 1967, and 1972 (54–60). In doing so, she concretely tied in with the affective-emotional experience of the human being, which both researchers had made strong, and which accordingly does not already express itself in discrete emotions, be it sadness, joy, anger, etc., but only in directional impulses, rhythms, and forms Cassirer ([1929] 1964, 88, 94, [1929] 1979, 75, 80; Whitehead 1927, 26–59, esp. 45, 2000, 101, 106; Sauer 2014b, 12–7, 2022). As Cassirer then elaborates in his late work, this kind of experience is retrieved precisely in the encounter with art. Thus, the content of art is “the verbally ineffable, yet not inexpressible law of vital experience, the pattern of affective and sentient being.” (Langer 1942, 209, [1942] 1965, 252). Understanding them demands familiarity with the “implicit” (rather than discursive or presentational) meaning of the works, which requires its own “non-discursive” forms of understanding (Langer 1942, 212–16, [1942] 1965, 256–60). The non-discursive forms of understanding correspond, as already indirectly discernible in Cassirer, to formal-abstract modes of image formation. According to Langer, the latter are based in music in the “tonal dynamic form” and in painting, sculpture, and poetry in the “play of lines, masses, colors, and substances.” Whereby the content of artistic expression itself, as Langer still suggested as a presumption in this early writing and made it a topic in Feeling and Form 1953, was the same in all arts as in music (Langer [1953] 1967, 103, 369, 372). Accordingly, Langer defined art as “the creation of forms symbolic of human feeling” (40). The concrete symbol-forming power lies in the fact that a “virtual” image of organic life emerges through the tensions and resolutions of the artistic means (206–07, 47–59, 372). It is based on the close relationship between organic (bodily, the present author) or mental processes (“vital forms”) and artistic (“artistic forms”). There must be an analogy between them. This is the basis of Langer’s (image-)act theory as presented in the double volume Mind. A Philosophy of Human Feeling, the first volume of which appeared in 1967 and the second in 1972 (Langer [1967] 1985, 199–253). Only in interaction, according to Langer, do they become symbolically (emotionally) significant. In the logical or dialectical pattern (“tension and resolution”) of the possible relations (“potential acts”), the meaning of appearance is formed as a “living form” or “illusion of bodily existence” (206).
However, to exclude this mode of action with reference to entertainment and ritual, as Langer assumes, because they pursue purposes, can hardly tenable. The assumption contradicts her own approach, according to which the processes of superordination and subordination of acts ("tensions and resolutions") not only constitute the behavior and thus the development of humans, but affect all areas of life. Thus, they can already be observed at the molecular level, which is controlled by physical and chemical processes. From a functional point of view, the foundation of developmental processes in dialectical acts thus enables the transfer of action-relevant information from one level to another. Understanding them consciously and translating them accordingly into media can therefore be regarded as the basis of human communication and development. Meanwhile, as mentioned in the introduction, recent studies in developmental psychology and neurosciences support this assumption (Stern 1985; 2010, [1985] 1992, [2010] 2011; Freedberg and Gallese 2007; Rizzolatti et al. 2013, Di Cesare et al. 2020).

Langer’s (image-)act understanding coincides in many respects with that of Cassirer expert John M. Krois (Sauer 2014b, 54–60). He, too, assumes that expressive understanding fundamentally constitutes human access to the world; it is physically bound (Krois 1987, 57, 85f.) and can accordingly be understood as a function of all higher symbolic forms (with symbolic conciseness). It is "the logical structure of experience" (47). Only late, however, in the last year of his life in 2010, does it become concrete for the researcher, in an engagement with enactivism (Freedberg and Gallese 2007; Krois [2011] 2010b, 237f.), that the understanding of an expressive meaning cannot only be related to the formation of will and thus to the desires and purposes pursued by it in a process of weighing ("evaluation process", Krois 1987, 155, 167, 102–05), but is related to the perceptual processes themselves. Thus, the theory of embodiment addressed by Krois in several recent essays gains contour (Bredekamp and Lauschke 2011; Sauer 2013). It is based on the assumption that both the image and the body (or the perception of humans, but also of animals and robots) are built on comparable principles or schemata. Accordingly, it can be summarized that there is an analogy between image schemata and body schemata: "In all of them, body schemata are built up from the same image schemata. These image schemata are dynamic, non-optical forms [...]" Experiencing and feeling them consciously (as "qualities", Krois [2011] 2010a, 231) through "expressive perception" (Krois 2011a, 270) distinguishes humans from animals.
Nevertheless, Krois asserts at various points—ultimately in contrast to his own statements—that although there are dynamic, affectively effective aspects in the image, which, as Cassirer and Langer also formulated, are based on abstract-formal principles, these are independent of the intentions of the producers and thus of the interpretation of the recipients (Krois, 2011a, 269; 2011b, 278; 2011c, 306).

The usurpatory character of pictural objects—the fact that they possess affective meanings independently of the artists’ intentions and the viewer’s deliberate interpretations—results from the fact that like the viewer, they too embody dynamic affective image schemas. (Krois 2010b [2011], 251)

Against this backdrop, it can be stated with Krois, that the image is not a communication and consequently has no relevance for action. This contradictory ambivalence also characterizes the art historical research of Horst Bredekamp, who founded Bildakt with Krois in 2008 (Bredekamp 2010, 51–6). In doing so, he disappointed the hopes of historians in particular, who, inspired by his lecture at the German Historians’ Conference in Constance in 2006 (2007, 289–309), hoped that the publication of Bildakt 2010 would provide operationalizable solutions for analyzing images in such a way that their agency would emerge (Jäger 2011; Sauer 2015b). Finally, it is the implicit presuppositions in Bredekamp’s approach, which can also be found in Gottfried Boehm’s work, that triggered fierce criticism because they would present images as independent actors and thus anthropomorphize them (Wiesing 2013, 78–107; Sauer 2014a).

Pipilotti Rist: I’m Not The Girl Who Misses Much, Video, 5’02, 1986

Contrary to this criticism, as it was formulated not entirely unjustifiably by Wiesing, both in Boehm’s and Bredekamp’s approach, against the background of the tradition of formal aesthetics within the art studies shown here and the cultural anthropological research as founded by Cassirer, an approach can be presented that is able to overcome the discrepancies. For if we assume that not only the perception of images, but already that of the world is based on non-discursive, formal-abstract aspects which can be taken in multimodally with all senses and, moreover, can be interpreted affectively-emotionally by us, then a conclusive concept emerges as to what extent unconscious, bodily processes can be seen as a prerequisite for how information and sensations translated by us into media can be understood. In particular, developmental psychologist
Daniel N. Stern has elaborated in research how fundamental this mode of perception is to social behavior and, moreover, to the perception of art (1985, 157–61, 2010, [1985] 1992, 228-30, [2010] 2011). However, unlike the interpretive processes in everyday perception, whose mechanisms of action go unnoticed, they become visible to us in the face of art. This reflects an assumption not only expressed by Stern (1985, 160-61, [1985] 1992, 225–230) and earlier by Werner (1926) 1959, 61–73, [1940] 1957, 59–72) but was also articulated in this way by Cassirer and Langer (Cassirer [1929] 1964, 88, [1929] 1979, 75, [1944] 1953, 189, [1944] 2007, 228; Langer 1942, 199–216, esp. 211, [1942] 1965, 241–260, esp. 254). Moreover, it seems self-evident that it is not in the interest of advertising and propaganda to reveal those processes through which they can unconsciously influence the sensations and thus the decisions of recipients (Sauer 2012a: 11–23, 194, 266–67, 283). It is noteworthy that in my view this contrasting interest is also reflected in recent film production in comparison to the computer game industry. While films such as Inception (USA 2010) or Gamer (USA 2009) are concerned with openly exploring the fictional possibilities of immersive processes, computer games aim to blur the boundaries between fiction and reality.

The following exemplary analysis of a video by the well-known Swiss video artist Pipilotti Rist is intended to show the extent to which the arts are not only concerned with revealing processes of perception and formation and their effects, but also with provoking reactions (Sauer 2012b). This is a video that the artist realized in 1986 during her formative years at the Basel Kunstgewerbeschule, now the Academy of Art and Design, and through which she became known at a stroke: I’m Not The Girl Who Misses Much (Fig. 1).

![Figure 1. Pipilotti Rist, I'm Not The Girl Who Misses Much, 1986. Video (video still), © Pipilotti Rist, Courtesy the artist, Hauser & Wirth and Luhring Augustine](image-url)
With the playful use of partly bizarre, partly funny effects and parallel video-technical procedures, the video artist not only caricatures a Beatles song, which also gave the video its title, but at the same time reveals the mechanisms of effect of the formation possibilities with moving images and thus also the modes of perception of the recipient. The following analysis pursues the opposite approach. It attempts to work out, via description, the specific formal-abstract aspects which, it is assumed, are affectively evaluated by us in parallel with our perception of the world and thus independently of culture. What influence this mode of perception has on the interpretation of the simultaneously recognizable elements and on the possibilities of active engagement with them must be shown, because this influence, according to the thesis, can be considered fundamental for the possibility of interaction and communication via media.

What the artist shows with the video is quickly told. In the tape, which lasts only five minutes, a young woman in a small black dress with open bosom dances and sings only blurred visible. Initially pixelated large, with a red filter and audible sequence soundtrack, the title of the video appears, enlarging and shrinking again. A sequence of events that we immediately follow and whose ambivalence we just as immediately transfer to the content of the text. It is this experience that shapes the interpretation of the meaning of the words, this can be emphasized here. The hectic movements and the face of a dancer behind the writing can be seen in fast forward and changing zoom settings, at first only very blurred. Then the shots switch to normal color. The lettering disappears and the dancer herself appears and becomes the subject. Still pixilated-blurred, color reduced to black and white with a slight red-green-blue shadow, her movements seem distorted puppy-choppy by the time-lapse. A short song lyric “I’m not the girl who is missing much” is repeated by her over and over again. The rhythm of the repetition and the time-lapse connects with the choppy movements. The performance style adds to this effect: the voice is breathless, high-pitched and squeaky. This performance experiences a new interruption as the medium itself becomes the subject through the flickering of the recording and the stripes in the scrolling. It becomes obvious, that the scene was shot with a still camera, and the dancer moves partly out of the frame or directly toward the camera. The performance is distorted or transformed by the use of a red filter (Figs. 2–3).
The image space is thus once again considerably flattened. In slow motion, as if from off-screen, the song lyrics are heard again, the dancer now lies below the field of vision, apparently on the floor, and slowly straightens up again. She takes the lead again and sings and dances even faster than before. Slow motion and fast motion, as well as the color change and the accompanying changing spatial impression, contrast violently with each other and intensify the effect. The performance, already bizarre at the beginning, takes on comic, slapstick traits in its urgency. Almost incantatory, the dancer sings the text over and over to herself and to us during her performance. All that happens through the staging of the formal elements and becomes affectively experienceable for us is here transferred to the scene and the text. Their context and meaning are suddenly called into question by the intensification of the effects. Then the film breaks down. Once again, the shot flickers and tear marks distort the scene (Figs. 4–5).
Questions about the meaning of what has been experienced and seen are not left out. But only with a deeper search for the references and background of the text it becomes clear that the video is directly related to a Beatles song written by John Lennon and Paul McCartney and published in the same year as the video, 1968. It talks about a girl and/or drugs, who lacks nothing sexually / intoxicatingly, but instead of “she’s” Rist says “I'm not the girl who misses much” from the song Happiness Is a Warm Gun. Ultimately, however, it is not so much the song and the girl as the staging techniques that irritate us. Both the girl and we with her are electrified by it: We experience the frenzy of the movements, dive into the red space, sink to the floor with her, become breathless, feel the rhythm and the turmoil. At the same time, we are able to reconcile these experiences with our own prior experiences and knowledge, so that we also simultaneously grasp the distortions or artificiality of the scene and the recording.
At a third level of experience, we relate all the information conveyed by the distortions to the dancer, her movements, and the song line she performed. In this way, we see how the artist’s creative actions or formal choices become essential to the production of meaning.

Understanding here proves to begin with an immediate experience and a “suffering”\(^4\) which “are at first a mere passivity, a being acted-upon rather than action” (Cassirer [1929] 1979, 75), and takes on concrete forms in comparison with one’s own prior experiences and prior knowledge. Accordingly, in retrospect, conclusions can be drawn about what the artist wants to tell us with the video: instead of having everything and thus missing nothing, as the girl pretends with the constant repetition of the song’s lyrics, she never seems to achieve this goal. Her attempts take on almost desperate-grotesque features. In the process, her performance never for a moment becomes concrete or clearer. On the contrary, what she wants to show herself and us threatens to slip away. She sings and fakes something for herself and for us. What she then shows us is the dancing and singing itself with bosom exposed and high-pitched squeaky voice. It does not gain any depth. In this respect, it is only a cliché or the image of a role that she wants to convincingly play for herself and for us. But just as the role-image does not work for the girl, the recipient also fails: instead of watching a seductive young girl herself/actualizing herself, this takes on clownish traits. It degenerates into a farce. This production, like others by the artist, completely eludes a possibly lustful, voyeuristic gaze or a possibly imitable role model function.
Conclusion

Perception in general, and thus also the perception of moving images produced with technical aids, is, as it was to be clarified against the background of cultural anthropological research and the exemplary analysis of the Rist video, not oriented solely factually to the recording of things and course of action, but allows itself to be affected by their formal-abstract modes of appearance. That is, it is oriented not only to what, but at the same time to the way something shows itself to us. In relation to moving images, these are the staging techniques and thus the respective settings with which the 18 or so images and their sequence are realized. Basically, it is they and thus the formally abstract modes of appearance that are consequently not only interpreted in terms of their possible nameable meanings, but at the same time captured in a bodily, affectively vital way. Decisive for the latter is, as a closer look reveals, how fast or slow, here how hectic or artificially slowed down a perceived movement turns out to be. Which direction it takes, whether it comes towards us or moves away. Equally significant are the sounds that accompany the movement, whether they appear suddenly and loudly like the rustling of the film reel or develop into a breathless, hectic, almost no longer intelligible chant, or whether they come from a distance, slowed down in tempo like the voice from off-screen. In addition, the shapes of the figures prove important, whether they elude us or constantly change their size and extension, as in the beginning of the song line and then in the case of the dancer through the blurred contours and abrupt movements. Finally, the constant change of space, alternating between real and color space and eventually culminating in the view of the tearing film reel, also contributes to the uncertainty. For the experience of the video, in this respect, it is the changes of form, the changes of direction, and the vehemence or degrees of intensity of the movements, as well as the changing speeds or time patterns of the movements and sounds, that determine our most immediate affective experience of what is happening.

It is thus both the movements of the dancer herself and the movements initiated by the artistic means, in this case the video staging techniques of the 1980s, that can be held responsible. This means that it is ultimately irrelevant for the experience itself whether it is afflicted by mundane or media-staged stimuli and by which techniques and materials it is to be accounted for. The stimulus value for the experience depends on the forms, degrees of intensity, and temporal patterns that trigger it, and on the immediacy with which it is experienced.
This is a connection that child psychologist Daniel N. Stern first worked out so clearly (Stern 1985, 47–68, esp. 52, [1985] 1992, 74–103, esp. 80; Sauer 2012a: 79–92, esp. 83–6). The remarkable thing about experience is that there is no difference between the real and the fictional world. Ultimately, it is one’s prior experiences and prior knowledge that can be seen as central to the degree of distancing and the corresponding conscious evaluation of what is experienced, depending on the degree to which they are admitted. That is, the less distancing processes are induced during enactment, for example by pauses, the higher the degree of immersion and the stronger the sense of one’s own body can fade, as recent empirical studies with reference to the experience of computer game worlds have shown (Weger and Lounghnan 2014).

Authors Biography

Martina Sauer runs the Institute of Image and Cultural Philosophy in Bühl (Baden), Germany: www.bildphilosophie.de. She is a senior editor of Art Style, Art & Culture International Magazine and a scientific advisor of the Society of Interdisciplinary Image Science and the German Society of Semiotics. She was a scientific associate in philosophy of art and culture, aesthetics and design in Basel, Zürich, Bremen, Witten as well as a scientific associate at Bauhaus-University Weimar. Currently, she holds seminars on the fundamentals of perception at Fresenius University of Applied Sciences, AMD Düsseldorf and Hamburg. Recently, she also initiated and realized a research project there together with Petra Leutner on “Affective Perception and Design” with the neuroscientists Giuseppe di Cesare and Giada Lombardi from Parma, Genua. For further publications see on: arthistoricm.net, researchgate.net, academia.edu, philpeople.org: Contact: msauer@bildphilosophie.de. ORCID: 0000-0002-4787-3189.
Notes

1. The study is the result of a close collaboration with the Research Group for Moving Image Studies in Northern Germany, Kiel, which in turn is closely associated with this present edition, following a joint conference. Recent research literature and some text passages have been added. In addition, writings by Ernst Cassirer, Susanne K. Langer, Alfred N. Whitehead, Heinz Werner, and Daniel N. Stern, originally cited from German editions, have been supplemented with text passages from English ones. In this issue this article appears in a revised and updated 2nd edition. Originally, it was published in German in the 2nd volume of the Yearbook of Moving Image Studies on the special topic of Image Bodies (Bildkörper). On the Relationship between Image Technologies and Embodiment, edited by Lars Christian Grabbe, Patrick Rupert Kruse, and Norbert M. Schmitz (Marburg: Büchner Verlag, 2016), 46–71.

2. In order to provide at least a glimpse of their research within the limited scope of this study, reference is made to some of their writings, most of which have been and continue to be highly influential in their field of study: on Warburg cf. Böhme (1997), on Panofsky ([1932–1964] 1984, [1939–1955] 1984) on Uexküll cf. Krois ([2007] 2011), on Werner ([1926] 1959, [1940] 1957, Sauer 2011), and on Grunow (1923).

3. In addition to the researchers already mentioned, the research approaches of Kurt Goldstein, a Frankfurt neurologist close to him (Krois [1999] 2011, 53–6, [2007] 2011, 188–90), the psychologist Ludwig Klages ([1913] 1950), the philosopher and psychologist Theodor Lipps (1899), and the philosopher, anthropologist, and sociologist Max Scheler ([1913] 1923) played a crucial role in his reflections on the theory of perception or the perception of expression. Cf. on the latter Cassirer's concrete references in the 3rd volume on the philosophy of symbolic forms ([1929] 1979, 58–91).

4. This reference to experiencing and “suffering”, goes back to the original text in German, which is much more concrete: “Ausdruck ist zunächst nichts anderes als ein Erleiden; ist weit mehr ein Ergriffenwerden als ein Ergreifen” (Cassirer [1929] 1964, 88).

5. This is initially followed by Stern’s joint research with Giacomo Rizzolatti’s research group in Parma (Rizzolatti et al. 2013; Di Cesare et al. 2020), which is currently being concretized and referring to art and design by Giada Lombardi and Giuseppe di Cesare (2022). Cf. their contribution in this special issue and the first joint contribution in collaboration with me for the next issue on the topic of Atmosphere and Mood (forthcoming 2023).
Bibliography


