

Article

The Perception of Depicted Motion

Livio Dobrez

College of Arts and Social Sciences, Australian National University, Retired, 9 Blair Street, Watson, Australian Capital Territory 2602, Australia; E-Mail: dobrezl@grapevine.com.au

Received: 27 August 2013; in revised form: 18 November 2013 / Accepted: 20 November 2013 / Published: 10 December 2013

Abstract: Everyone knows that you can read a galloping horse in a still image as galloping. This paper asks how it is that we perceive motion in pictures. It considers perception of real motion in point-light experiments and the perception of motion in stills via the work of various psychologists, in the course of which it raises theoretical questions about the nature of visual perception. It then offers a detailed examination of knowledge regarding neural substrates for both real and depicted motion perception. Finally, it combines psychological and neurophysiological perspectives with phenomenologically-oriented observation of pictures, discussing both frontoparallel motion and motion in depth (in particular the phenomenon of “looming”) in terms of two kinds of depictions, the “narrative” and the “performative”. Examples are drawn from all kinds of pictures, but focus is on world rock art, whose time depth is especially amenable to the universalist approach adopted by the paper.

Keywords: visual perception; depicted motion; rock art; perceptual psychology; neurophysiology; phenomenology; point-light experiments; representational momentum; intention; causality

1. Real, Apparent, Imagined and Depicted Motion

How do we perceive motion? How do we perceive motion in a still picture? Some of the experimenters discussed below talk about motion perception, some refer to it as event perception or, on occasion, perception of action. A person, for example, waving an arm, is of course *moving*, but also performing an *action*—which constitutes an *event*: X waving. So I intend to use these terms interchangeably, though I will eventually make a distinction between action and interaction.

There are various kinds of perceived motion. We see *real* movement. We see what is referred to as *apparent* movement, that is, an illusion of movement. If we visually track a moving object we register a background shift in the opposite direction (the Filehne Illusion). Illusion is exploited in Op Art, where apparent movement is evoked by geometric patterns, wavy lines *etc.*, as in the abstracts of Victor Vasarely or Bridget Riley. Equally, it operates in animated cartoons and in the movies, where we read a rate of 24 frames-per-second as motion (the silent movie 16 fps looks jerky). We are able to *imagine* movement: we may mentally visualize galloping horses or (if we are perceptual psychologists) mentally visualize a rotating cube (otherwise known as performing a “rotation transformation”). Finally there is *depicted* movement. I shall not argue that we *infer* motion in a still picture of a moving object. It is not a question of *knowing* that the horse in the picture is galloping, but rather of *seeing* it as galloping. So my concern is with perception, not inference, though I realize that some take the view that perception is itself a process of inference. Specifically, my concern is with perception of depicted motion—in any kind of picture or representation, but especially in rock art which has time depth unmatched by other forms of representation. Focus on perception means that I shall not consider culture-specific, for example, symbolic interpretations, but perceptual universals which, as far as we know, have not changed in the last several hundred thousand years, since the change in brain size between *Homo erectus* and *neanderthalensis*. In fact, as experiments, which are mostly with monkeys, show, our visual system has been relatively stable over the 20 million years or so that separate us from monkeys. I return briefly to this matter of visual constancy in section 19 below but at this point it suffices to make the confident assumption that where we see movement in a picture, so did humans in the oldest relevant art, some 30,000-plus years ago.

I shall consider depicted motion in connection with perceptual experiments and neurophysiology, before offering phenomenological observations of my own, many of which should be testable.

To begin with, however: some illustrations which bear on issues raised by this paper. In rock art especially (but not only in rock art) there are many examples of depicted motion likely to be judged borderline.

A composition from the Palaeolithic site at Niaux in the French Pyrenees (Salon Noir, Panel 6: Figure 1) features formally unrelated animal figures which I would term “juxtaposed” and two figures at bottom right which might suggest something more like a necessary “association” (for definitions of terms, see Dobrez 2011 [1]).

The two bison face each other but do they do so in a purely formal sense or in the sense of an *activity*? So is this “facing” a case of depicted motion, though no motion is apparent? And if no motion is apparent, might motion nonetheless have been intended? I shall consider the problematical matter of “intention” later in this paper, but in terms of what is apparent, that is, available to perception (or “reception” in the language of hermeneutics) I am inclined to say it is simply a suggestive formal association. What about this association of Australian canines and macropod, perhaps a wallaby or kangaroo (Figure 2)?

Is the association stronger than in the case of Niaux and, if so, does it constitute a picture of an event, in short, are the dogs “chasing” the kangaroo? Possibly so—just. Here is an entirely different kind of example: the falling-man sign certainly depicts movement, but is not relevant to my argument here, the reason being that it is not a picture of a man falling, but a message to the effect that we must take care not to slip on the wet floor (Figure 3).

Figure 1. Niaux (model), Parc de la Préhistoire, Tarascon-sur-Ariège, France.



Figure 2. Yankee Hat, Australian Capital Territory.



Figure 3. Workplace sign.



Putting it another way, it is indeed a picture of a man falling, but available cultural information allows us to make a clear inference as to its meaning. We cannot *see* this meaning, but we can *know* it. And the meaning is not “falling man” but “don’t slip”. For this reason, the picture (cultural information directs us to the term “sign”) falls outside the scope of my investigation. A sign (Figure 4) of gravel hitting a vehicle also depicts movement, but it is not a picture of gravel hitting a car. Rather it warns of

gravel on the road. I refer to signs of this sort as “symbolic”, though not in the Peirce sense of non-iconic, and the reason they do not concern me here is that symbolism may be added to any representation whatsoever and that it depends entirely on culture-specific information: to read it you need to know the code. Whereas the present argument, as explained, deals with perceptual universals, characteristics not limited to time and place. Presumably, though, the famous Boar Rock images in central India (Figure 5), whatever their symbolic meaning (which is unknown), are indeed a picture of an event—as are those from the “battle” scene at Writing-On-Stone, the celebrated site in southern Alberta (Figure 6: reduced from ink tracing on polyethylene by J. Keyser, 1976):

Figure 4. Road sign.



Figure 5. Boar Rock, Bhimbetka, Madhya Pradesh, India.



Figure 6. “Battle” scene, Writing-On-Stone, Alberta, Canada. Courtesy James Keyser.



In which case, what is it about these last examples that allows us to perceive them as depictions of motion, presumably in somewhat different ways? Finally, does the great panel from Dinwoody Lake, Wyoming (Figure 7) have any relevance to the question of depicted motion?

Does this depict motion at all? If it does, it must relate to a quite different sort of motion than that in other examples just cited. I shall return to questions raised by the above representations in due course, following discussion of experimental data available for the argument I shall be putting forward.

Figure 7. Dinwoody Lake, Wyoming.



2. Perception of Motion: Point-Light Displays

We may perhaps think that when we see a moving object what we see is its form, that is, that we see *things*—which happen to be in motion. However, as that highly original psychologist, Gibson (1950, 1966, 1979 [2–4]) has insisted, in real perceptual situations everything is undergoing transformations, that is, everything is in flux. Not only does the visual object move—so does the perceiving subject, whose eyes, head and indeed entire body is likely to be in more or less constant activity. In such a situation can we simply say we see things—moving? Might it not be more precise to say that evolution has geared our visual system to register motion itself? In Gibson’s words: “optical rest is a special case of optical motion, not the other way around. The eye developed to register change and transformation” (1979 [4], p. 293).

A number of studies suggest this is so: we see the movement of things before we register their form (Mather and West 1993, Dittrich 1993 [5,6]). The pioneer in these studies was Johansson (1973, 1976 [7,8]). Johansson devised experiments with so-called point-light displays, in which we attach lights to the joints of a real human body and have the person move in the dark while observers register whatever it is that is going on. The results are astonishing: observers inevitably report perception not of light points but of coherently-moving figures. As few as five to ten points adequately register what is termed “biological motion”: observers accurately distinguish not only walking, running,

dancing, *etc.*, but a limping motion, a “tired” motion, a particular individual’s motion. Clearly reading “form-from-motion” is going to have considerable survival value: it takes less time to register movement than to identify a specific form in movement. The key thing to note being that when point-light figures stop moving, identification fails. Observers no longer see a human figure, only a jumble of lights.

Now for Johansson such experiments suggest perceptual immediacy rather than application of learned motion patterns. We do not *infer* the form from its motion; we *see* the form and we see it in seeing the motion. This is a Gibsonian conclusion which today might be taken as supporting the thesis that we are hardwired to spot movement, and to read it as a given form in very quick time. Dittrich [6], who extends Johansson’s experiments with biological motion, additionally argues that motion is perceived at a complex organizational level of activity. Stressing the cognitive capacities of perception, he ties the idea of an action-category to the work of Rosch *et al.* (1976 [9]) on categorization (for comment on which see Dobrez and Dobrez 2013 [10]). While it seems to be the case that we pick up high levels of meaning in the simple act of seeing, this need not be referred back to the “learning” model of perception, which requires the intervention of inferential processes. Rather we read sophisticated visual meaning directly, something readily explicable in evolutionary terms as neural corner-cutting. Walk (1984 [11]) takes up this Gibsonian option of “direct” perception, referring to the Johansson phenomenon somewhat poetically as the “sound of silence” in visual perception, that is, the seeing of a form in its “absence”, or as a presence indicated only by its motion.

At any rate the general thrust of point-light experiments is well established. The principle of form-from-motion applies when lights are placed *between* joints as well as at joints (Dittrich [6]). Remarkably, it applies in a point-light situation in which an otherwise invisible figure lifts an otherwise invisible object—and observers actually register the *weight* of the box (Runeson and Frykholm 1981 [12])! Bertenthal *et al.* (1984 [13]) have shown that three- to five-month-old infants discern figural coherence in a walking human even when the display is inverted—but not if the figure remains motionless. This is not very different from Dittrich [6], whose inverted moving figures are still seen as moving even when the exact type of movement is unclear. Mather and West [5] combined the point-light method with Muybridge’s celebrated photographs of figures in arrested motion. As I read it, their work further strengthens the view that biological motion (in this case with animals) is hardwired (see also Shepard 1984 [14]). Again, their observers identified point-light features as given animals when movement occurred and not with static displays. All this underlines the Bertenthal *et al.* comment: “one might suppose that the extraction of form from a continuously changing object would be more difficult than the extraction of form from a stationary object” ([13], p. 213). It is not so—rather the reverse. We really do appear to see motion first, and for excellent evolutionary reasons, since in many situations having to see the object’s form *before* its motion might just be seeing-too-late! Perception is always required to anticipate; you have to see *more* than there is to see. Put concisely: only seeing-more is sufficient for seeing. And this suggests a further observation: that the human brain, of which much will be said below, is a near-enough instrument. Evolution has to err on the side of caution, and if you see “something like” X, you will assume it *is* X, at least provisionally. What perception of motion does is to signal as much what *might be* as what is. It should come as no surprise that in point-light experiments “near enough” is the procedural rule—and, equally, that it turns out to be “good enough”. As Mather and West [5] put it: “even drastic mislocation of dots does not preclude

correct interpretation” (p. 762). Seeing is “near enough” because it has to be quick. But “near enough” has to be “good enough” to ensure survival.

3. Mental Representations of Motion

A series of experiments takes up the issue of perceived motion where point-light experiments leave off. In particular Freyd (1983 [15]) makes a move which is critical for the present argument: she does this by considering *depicted* motion, though depiction as such is not her concern. So in referring to her work I realize I am taking it very much out of context, or at least well beyond its immediate context. However, I make no apology for this, because what Freyd does is directly relevant to the matter of explaining the perception of motion in a static image. For present purposes, I simply want to emphasize her view that we see movement “independent of whether the stimuli are dynamic or static” (p. 575), that is, regardless of whether the object we observe is moving or static. I shall unpack this statement in due course but first would like to go into the issue of “mental representations”. When Freyd formulates her experiments in terms of mental representations of movement she is following many contemporary perceptual psychologists, but going against the redoubtable Gibson. Gibson did not believe perceptual theory required the idea of mental representations. Seeing was “direct”, it needed no mediating processes, in particular conceptual or inferential ones. But Freyd’s argument is not for cognitive or inferential mediation (as proposed for example by Pylyshyn (1981 [16])). Rather it falls on the “analogue” side of the analogue/propositional debate in the psychology of perception. Thus her “mental representations” of perceived motion are not concepts but percepts. The case for a middle way between the Gibsonian position and those who espouse a need for mediation by way of mental images is tactfully and convincingly put by Shepard [14] in a paper which originated in the Gibson Memorial Lecture he gave at Cornell in 1983.

Shepard sought to reconcile Gibsonians and their opponents by accepting Gibson’s contention that we see “directly”, that is, without mediation, all necessary information being in the optic array available to us from our environment—but all this only in favourable visual conditions. Where direct perception does not suffice, say in situations of poor light, partial occlusion of the object we are observing, not to mention situations of brain damage in the observer, degraded visual information requires the assistance of mental representations. These representations naturally come into play when we “see” what is not actually present, for example when we are planning ahead or anticipating, or when we are imagining or remembering or dreaming or hallucinating. The key thing being that these phenomena, not least dreaming or hallucinating, are not chaotic but follow laid down perceptual patterns. Like circadian rhythms which continue to operate even when their light stimulus is removed, basic perceptual patterns have become *internalized*, that is, mediated, as a result of evolution, by neural mechanisms. In situations of degraded visual information these will kick in and make up the deficit. In this way the brain adjusts to injuries to the visual system or helps us to see partly occluded objects, *etc.* Shepard thinks of it as a system of “internalized constraints”, that is, real-world perceptual constraints or structures currently operative in our neural systems, in short, hardwired. The relevant point here being that such constraints take the form of mental representations which structure our seeing. Whether or not Shepard’s gloss on Gibson, with the addition of some element of mental representation, is or

should be on the way to acceptance by Gibsonians (e.g., Neisser 1993 [17]) is a question that may be left open in the present paper.

Shepard relates the above to an idea of “resonance”, something which may, in the context of rock art analysis, explain a variety of visual phenomena, not least self-generated imagery presumed to feature in “phosphenes”. In particular, for Shepard, it gives an insight into the hierarchical nature of the visual system, with high-level operations which are complexly integrated, and which enable fast switching between operations without as it were constantly restarting from scratch with elementary building blocks. My interest here is that this might provide a structural explanation for visual switches discussed below. Additionally, Shepard’s argument would make sense of what I termed the “near enough” principle. When visual information is ambiguous or incomplete, the system will tune in to the nearest appropriate neural station. And of course this would ultimately shed light on the known intimate neural connection between seeing, imagining and remembering—as well as the (not so well known) connection between seeing and seeing *depictions*, to which I now turn.

4. Representational Momentum

As stated above, Freyd’s interest in mental representations of motion is with reference to seeing objects that may be in motion—or *not*—and it is the latter which is the focus of experiments of major relevance to the present argument. A simple experiment (Freyd [15], Figure 8) consists of showing observers a photo of a man jumping down from a wall (photo 1: top), then removing the image and, after a specified interval, showing a second photo of the same event (photo 2: bottom).

Figure 8. Implicit Motion: before (**top**), after (**bottom**). Courtesy Jennifer Freyd.

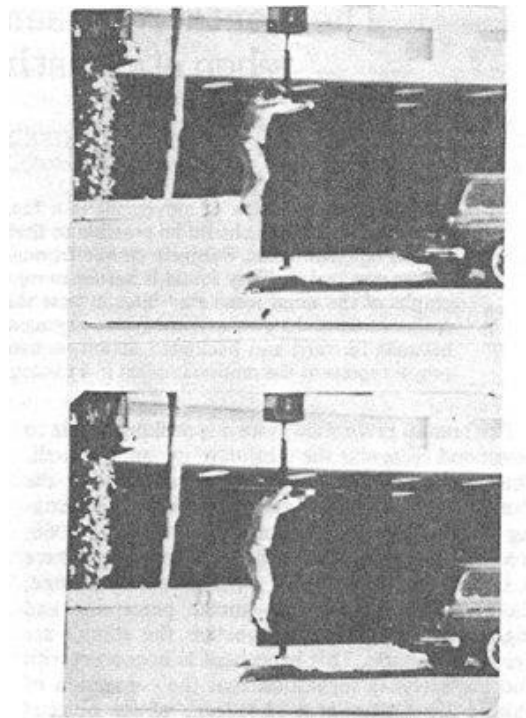


Figure 1. A before (above) and after (below) pair from the action sequence showing a person jumping off a wall.

These Figure 8 images are static depictions of motion, with the second picture showing the man as having jumped a little beyond the position of the first. Freyd asked her subjects the question: are the two pictures “same” or “different”? Observers overwhelmingly registered “same”. But the pictures were *not* the same. In fact the second still image had the man positioned further along the jumping trajectory than the first. Freyd’s conclusion was that we see photo 1, retain it in short-term memory, then, when exposed to photo 2, compare our memory-image of photo 1 with photo 2. But our memory or “mental representation” has *shifted* the movement in photo 1 along a bit, before we arrive at photo 2—which is why the second picture is deemed the same as the first. In other words, memory overshoots the mark, which signifies that we actually saw that still figure as being in motion, and by a measurable overstated distance. Freyd called this “implicit motion”, motion perception cued by a still or frozen-action shot. It should be emphasized that there was no question of a visual illusion, simply of reading motion in a *picture* of motion. Of course we know perfectly well that we can read depicted movement, but it was no small thing to find a way of quantifying the phenomenon.

Freyd and Finke (1984 [18]) went on to argue for an identical visual effect in connection with the direction or path of movement (Figure 9).

Figure 9. Representational Momentum. Courtesy Jennifer Freyd.

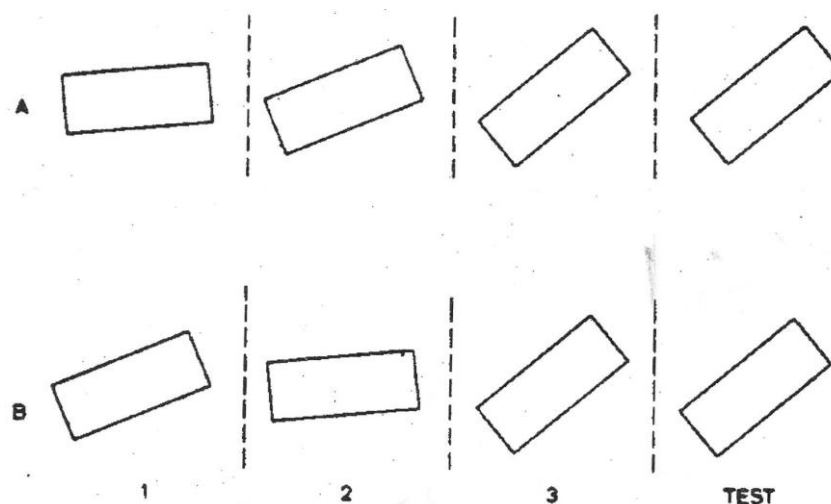


Figure 1. Examples of display sequences. (Panel A = Experiment 1 and Panel B = Experiment 2. The three stimulus presentations plus the test pattern, which were presented sequentially in the experiments, are shown here from left to right. Test patterns shown on the far right are for the *same* trials, in which the test-pattern orientation matched that of the third pattern in the sequence. Distractors consisted of identical patterns rotated 0.1 rad, approximately 6°, clockwise or counterclockwise from the third pattern.)

They exhibited three rectangles at various stages of rotation, consistent or otherwise with respect to direction, then a fourth which matched or did not match the third. Critically, in some cases the fourth slightly overshot the third. When this occurred in combination with consistent rotation in a given

direction, the fourth rectangle was judged “same” as the third, that is, subjects recalled the third as having gone further than was the case. The effect was destroyed by inconsistency in the direction of rotation, as in panel B, but *not* panel A, which is consistently anti-clockwise (see Figure 9 above). Freyd and Finke dubbed this tendency of subjects to shift their memory of the third rectangle in a forward direction “representational momentum” (RM). Clearly “implicit motion” and RM are related in that both entail visual displacement. If, following Shepard, we take the view that mental representations operate as internalized laws of physical motion, then the “implicit motion” of the man jumping from a wall would be read as an internalization of Newton’s law of gravity and the RM of rotating rectangles as an internalization of Newtonian momentum, the tendency of things in motion to continue that way. Freyd and her colleagues have taken a less than literal approach to the phenomenon, though as I understand it, they support the internalization thesis. Perhaps a little confusingly, others have used the term RM as a general description of visual displacement. In what follows I shall generally speak of RM (an attractive shorthand), but refer to the jumping man as exhibiting “implicit motion”. In the end, though, it is vital to understand that while Freyd is interested in Newtonian momentum, I am concerned with perception of motion in a still picture. Freyd shows that such motion is actually measurable, and for that she requires her sequences of images which essentially reveal *remembered* motion. I am interested in *perceived* motion, and what these sequences demonstrate to me is that remembered motion presupposes perceived motion. You *recall* the motion of the penultimate item in the series precisely because you have already *seen* it. Put more pithily: you cannot remember what you have not seen. This is the context in which Freyd’s experiments become critically important for my own argument.

Given the above account, it comes as no surprise that Freyd (1987 [19]) specifically put the case that some or most mental representations (conceivably, concepts as well as percepts) are dynamic, that is, that they include a temporal dimension. Now Gibson’s perceptual theory famously allowed for temporality and Gibson understandably drew fire by stating that we do not just see what is before us but, as it were, around the corner—an idea with Gestalt roots in Koffka but which became central for Gibson. We see “around the corner” because all our seeing is of a piece, that is, we see continuously rather than discretely, such that perception incorporates time. Less polemically, Shepard concluded his revision of Gibson by noting that what allows for perception of past and future, as well as present, events is precisely the fact of mental representations—which structure the present (where required), store past events and, finally, anticipate futural ones. Freyd’s argument borrows from Gibson in proposing that, from the standpoint of perception, a static object is simply a special case of an object in motion—though of course she translates this into the language of mental representations. Thus “static representations, if they exist, are special cases of dynamic representations” ([19], p. 436). Dynamic representations include temporality in an intrinsic structural way. This means that they exhibit the characteristics of time, namely continuity and one-way directionality (the so-called “arrow of time”). It also means that temporality is a necessary element of the representation: it is as it were on the “inside”, not, as postulated by other image-based theories, on the “outside”. So at this point the case put is a more radical one than Shepard’s. It suggests perception of a given motion or event as including its previous state as well as its (“implicit”) future one—*within* the representation of that motion or event.

Freyd’s experiments have been repeated and extended by others. With some reference to RM, Schütz-Bosbach and Prinz (2007 [20]) investigated anticipation under the rubric of “prospective

coding”. Thornton and Hayes (2004 [21]), using video, showed RM operative with complex visual sequences involving anticipated action, and Blätter *et al.* (2012 [22]) showed it as operative in situations of real movement. To my knowledge nothing has been done in the way of connecting Freyd’s insights with analysis of images in art (with the exception of a passing reference cited below). Freyd herself refers to a chapter by Friedman and Stevenson in Hagen’s *The Perception of Pictures* (1980 [23]), somewhat in desperation, perhaps, since the chapter (not the book) seems to me confused. I realize that, as Hagen explains in her role as editor, Friedman and Stevenson are to be taken in the context of the book’s focus on projective perspectival representation. Even so, their treatment of what they call pictorial movement “indicators” sheds more darkness than light on the subject. I return to such indicators in due course.

5. Percept and Concept

Now Freyd judged that, given the completely automatic nature of the RM response, namely the fact that we cannot control our visual tendency to overshoot the mark by seeing movement in a frozen-action photograph, the operation was perceptual rather than conceptual. Rapid mandatory response meant that the mechanism had been internalized (broadly in the manner proposed by Shepard) and so was impermeable to properly cognitive factors. This view counters Pylyshyn’s propositional or inferential case for perception as affected by “tacit knowledge”, that is, as “cognitively penetrable” (Pylyshyn [16])—in which scenario mental representations of the sort under discussion here would be influenced by our thinking on the subject. Of course RM, understood as an automatic response, would pre-empt cognitive input. It happens that a number of experimenters (not necessarily entirely in step with Pylyshyn) have queried Freyd’s interpretation, putting the case that RM *is* affected by knowledge of the situation. The debate has been summarized in detail by Hubbard (2010 [24]). Hubbard and Bharucha (1988 [25]) had observers view a circle (liable to be read as a “ball”) moving towards the edge of a frame. Sometimes the circle “bounced” off the frame after “collision” with it, sometimes it went “through” the frame, that is, crossed the line, and continued its linear motion. The point of the experiment was to cue the observers to *expect* either collision/bounce or breakthrough. In each case the Freyd RM rule obtained: subjects expecting a bounce underestimated and overestimated motion by registering *slowdown* just before and at collision and *acceleration* in the bounce-back; subjects expecting a breakthrough registered continuation of linear trajectory, motion going further than was actually the case, but in this case in a straight line. In other words expectation read RM differentially in bounce or breakthrough on the basis of a specific visual cue given as a prelude to the experiment itself. The authors took this as evidence of a high-level cognition operating in RM: the visual was apparently permeable to semantic factors. They suggested that a further experiment using strictly semantic, that is, verbal cues rather than their visual ones would strengthen their conclusion.

Such an experiment was conducted by Reed and Vinson (1996 [26]) who did not *show* but *told* their subjects what to expect, then measured the RM shift. Their conclusion was that RM could not be innate, as Freyd supposed, since subjects viewing an upward-pointing object varyingly positioned in a series of stills (as in Freyd’s experiments) exhibited a *greater* RM response when told it was a “rocket” and a *lesser* when told it was a “steeple”. The object in fact was pretty indeterminate, but it was more seen as “going up” if read as a rocket and less, in spite of its pointed shape, if read as a steeple. Reed

and Vinson thought this result was more in line with Shepard's "internalized constraints" than was Freyd's conclusion, though of course their experiment entirely backed Freyd's idea of RM. They suggested that we *learn* the typical movements of objects and internalize this information so that its visual operation becomes automatic. This amounts to a particular interpretation of the idea of "internalized constraints" in relation to memory. Freyd allows for short-term memory as a factor in her work, since her RM is precisely defined as distorted memory: we judge more shift than we see—on the basis of our *recall* of an earlier image. On the other hand Reed and Vinson's appeal to knowledge of "typical" motion requires long-term memory intervention in the perceptual process. Such intervention is feasible when, for example, we recognize an object on the basis of a previous perception (for example registering rockets and steeples), but I cannot readily see it operating in the case of motion perception. Putting it another way: it is not evident that facts learned through individual experience are going to generate visually automatic responses. Likewise I am not sure how we are to reconcile individual long-term memory of the Reed and Vinson sort with evolutionary adaptation of the sort envisaged by Shepard: it would have to be very long-term indeed, so much so as, once again, to depart from the conceptual learning model.

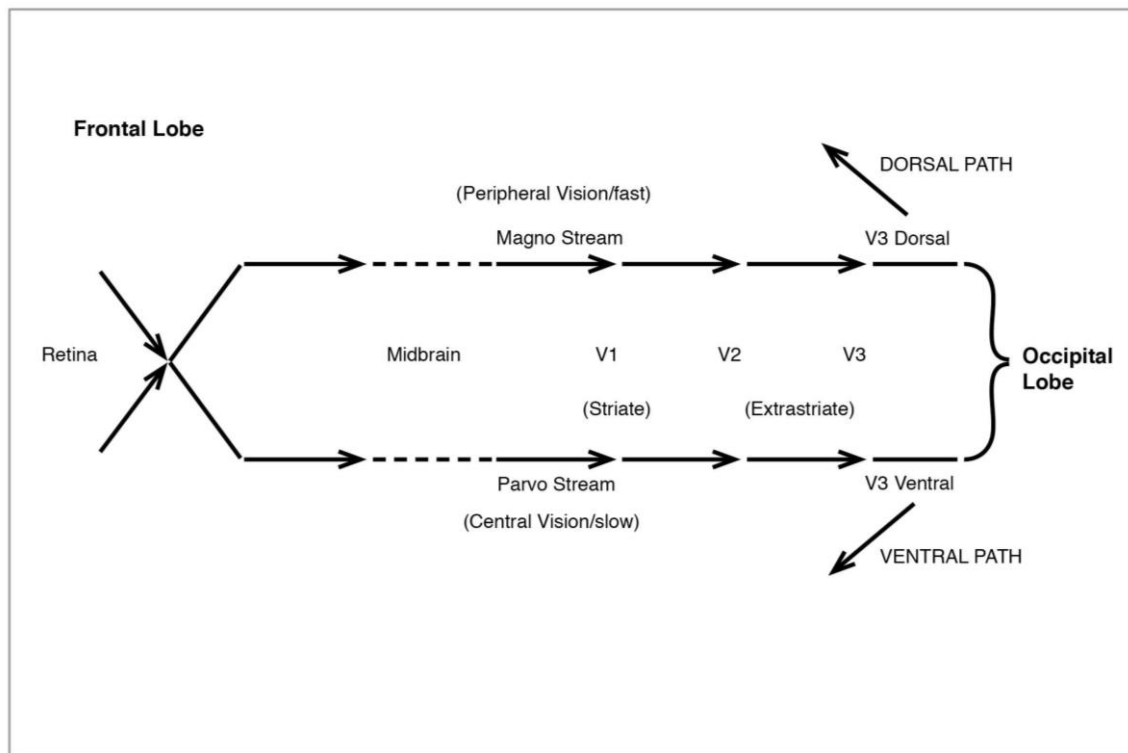
Reed and Vinson call theirs an "interactive" model, combining Freyd's stress on RM as automatic with Pylyshyn's "cognitive penetrability". But the interaction is one of percept and concept—which is precisely what one understands by "cognitive penetrability". So the argument would seem to return to Pylyshyn and the primacy of concept-based or propositional approaches. Rightly or wrongly I find the model an awkward one. Taken to a logical conclusion it would postulate a situation in which we see a picture of motion, then *infer* or deduce what happened earlier and what will happen next on the basis of acquired knowledge of how things usually behave. Even for a visual system capable of very rapid processing that would seem a cumbersome route. It is true that the human brain is not required to operate by the laws of Aristotelian logic. Likewise Freyd's own experiments reveal RM responses that are sometimes more, sometimes less "robust". Nonetheless the automatic nature of the RM response is immensely compelling and in the light of that alone it may be as well to seek for its explanation in perceptual rather than conceptual terms—without ignoring the apparent effect of verbal cues in Reed and Vinson's experiments, but also without following Reed and Vinson to their conclusion, namely the primacy of a propositional approach. For a start it is not obvious that the Hubbard and Bharucha experiments pose an insurmountable problem for Freyd's thesis: after all, they operate within the sphere of the perceptual. What they show is that perceptual shifts or switches are perfectly possible, which comes as no surprise. But a shift from percept A ("bounce") to percept B ("breakthrough") is not equivalent to an interaction between percept and concept—in which the latter determines perceptual outcome. It might simply be said that to cue experimental subjects, certainly visually, as with Hubbard and Bharucha, or perhaps even verbally, as with Reed and Vinson, is to prompt them to "throw a switch" in a manner analogous to figure/ground: "see X as *either* this *or* that". The switch need not be envisaged as a cognitive intrusion. In which context it might be reasonable to judge that RM is *not* the outcome of learned behaviour. Of course it may well be affected by expectation of particular perceptual situations with their appropriate internalized constraints. However, once a given reading of the situation is activated, once, presumably, a given neural path is activated, perception has no choice: internalized constraints dictate RM and we really do see motion in a still.

Conceivably an argument along these lines applies to a comical example used by Senior and Foley (2006 [27]) in an online report. The authors illustrate their point with a picture by David Hockney (“The Picture Emphasizing Stillness”) in which two apparently unaware victims chat while a leopard leaps towards them. Naturally we read “leaping”, that is, motion. But there is small print between leopard and victims which says “they are perfectly safe, this is a still”. At which point we “freeze” the leopard, that is, *undo* “leaping”. Senior and Foley think this is a case of cognitive input affecting the mental representation of motion, as in the Reed and Vinson experiment.

However, it is not necessary to accept the conclusion. It goes without saying that we are able both to see motion in a picture and to freeze it. In fact, we are capable, if we try, to read motion *backwards*. But the point is that some readings are *easier* than others, that is, more in line with internalized perceptual constraints. It is easier to read the leopard as leaping than to freeze it, and easier to read it as leaping in the direction of its orientation than as leaping backwards, as Ninjas do in Japanese films. We see forward motion—where motion is depicted—more easily than any other because evolution has internalized momentum and not because we learn by experience that motion is generally forward once it has started that way. Of course experience does not contradict laid down, that is, innate perceptual structures. How could it? These structures have been laid down in evolutionary time precisely through visual interaction with our environment. At the same time, *depicted* movement is not *actual* movement and we see it as such. Consequently we are perfectly free to halt it in its tracks, that is, to freeze it. Hockney plays with this visual option, though the fact remains that it is easier to see the leopard as leaping than not. Or, again, that it is possible to switch from motion to still in a (still) picture, in a way not possible with actual motion perception. The switch is presumably related to the capacity for *attentional* switches in real-life situations, noted by Gibson (1971 [28]) and explained along Information Theory lines. In connection with the percept/concept debate referred to above, one thing worth bearing in mind might be that, if I may so put it, perception is older than cognition, at least if we define cognition not as “meaning implicit in perception” (what Gibson termed “affordances”) but as “meaning that is tied up with language and thought”, that is, propositional meaning. In terms of this latter definition of cognition we can say we have been *seeing* much longer than we have been *knowing*. So it probably makes sense at least to begin by prioritizing the percept. But the phenomenon of an attentional switch, suitably applied, might allow for the conclusions of more semantically-oriented researchers, for example Hubbard and Bharucha who interpret RM as a high-level neural operation—though not Kerzel (2002 [29]) who interprets it as low-level (unless, as Hubbard [24] tells us, he has shifted his position!).

6. Neural Pathways for the Processing of Visual Motion

We have reviewed psychology-based research on the perception of movement with the conclusion that in all likelihood we are geared to see movement over and above static objects, and that this fact extends to depicted movement as “implicit” motion or as RM. But what might constitute a neural basis for all this? (Figure 10).

Figure 10. The Visual System: initial processing.

The processing of our visual experience of the world begins with light registering on the retina of the eye. Light is initially processed by a variety of cells, ganglion, bipolar, horizontal and amacrine, as well as the rod and cone photoreceptors. Of these the ganglion cells, which come in two types, large and small, have special significance for what follows. Signals from the eye pass through the midbrain thalamus region en route to the visual cortex at the back of the head, and they do so in two streams, the “magno” and “parvo”, which have been linked to the original retinal division of large and small ganglion cells. When they reach the visual area proper at V1 in the occipital lobe, they are still operating as two streams. Much of the work on V1 and the idea of magno/parvo signals has been done by Hubel and Wiesel (1959, 1962, 1968 [30–32]) and is described in a book designed for the general public by Livingstone (2002 [33]). Hubel and Wiesel exposed cats to light beams and noted the effect on individual neurons in the early part of the visual system (V1). They found that light stimulus had no effect if it was diffuse, but registered very specifically when passed through a narrow slit. Moreover the angle of the slit (vertical, horizontal, *etc.*) registered in specific neurons geared to respond to that particular angle and no other. It seems that diffuse light was neutralized by an inhibitory effect in the cell, but that focussed stimulus identified individual neurons responsive to a given orientation. This meant that early processing of light is concerned with distinguishing visual discontinuities. In fact there was a cell hierarchy in V1 which began with orientation, eventually extending to the contours of objects, in short, generating an initial reading of shapes. At the same time movement registered very strongly, unsurprisingly, since survival requires the earliest possible signal that something is on the move in our visual field.

After early processing in V1 (also known as the “striate” cortex because of its characteristic stripes), signals go to the “extrastriate” areas of V2 and V3, in a magno/parvo division which becomes

especially marked at V3 (sometimes divided into V3d and V3v), then diverge entirely (Figures 11 and 12).

Figure 11. The Visual System: ventral pathway.

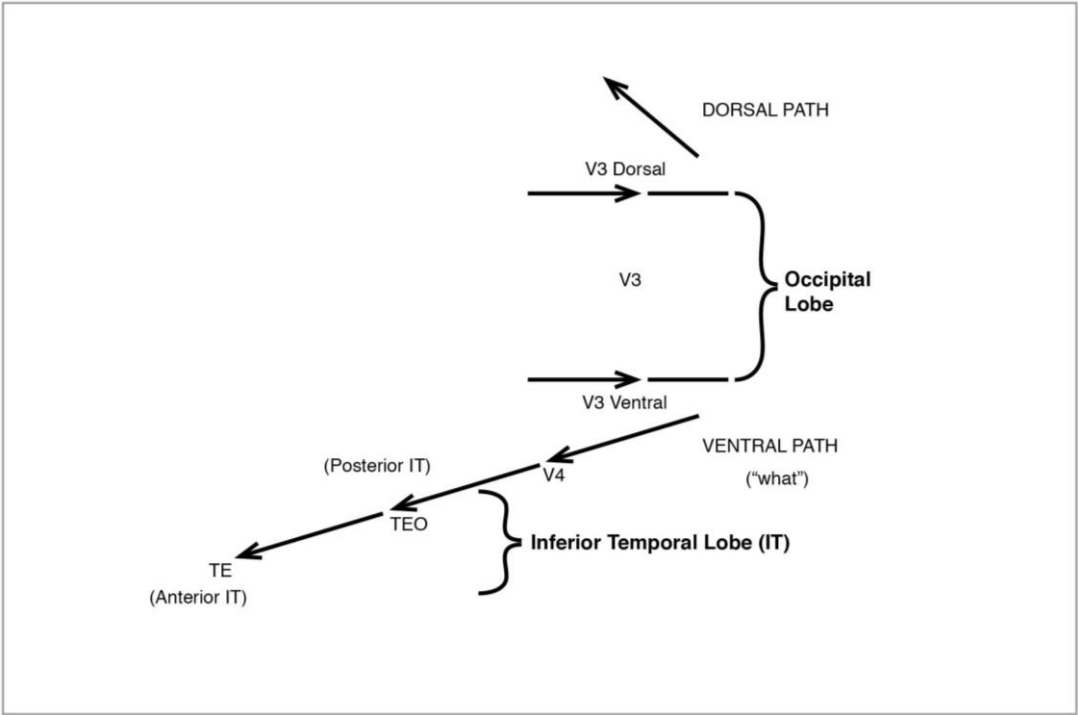
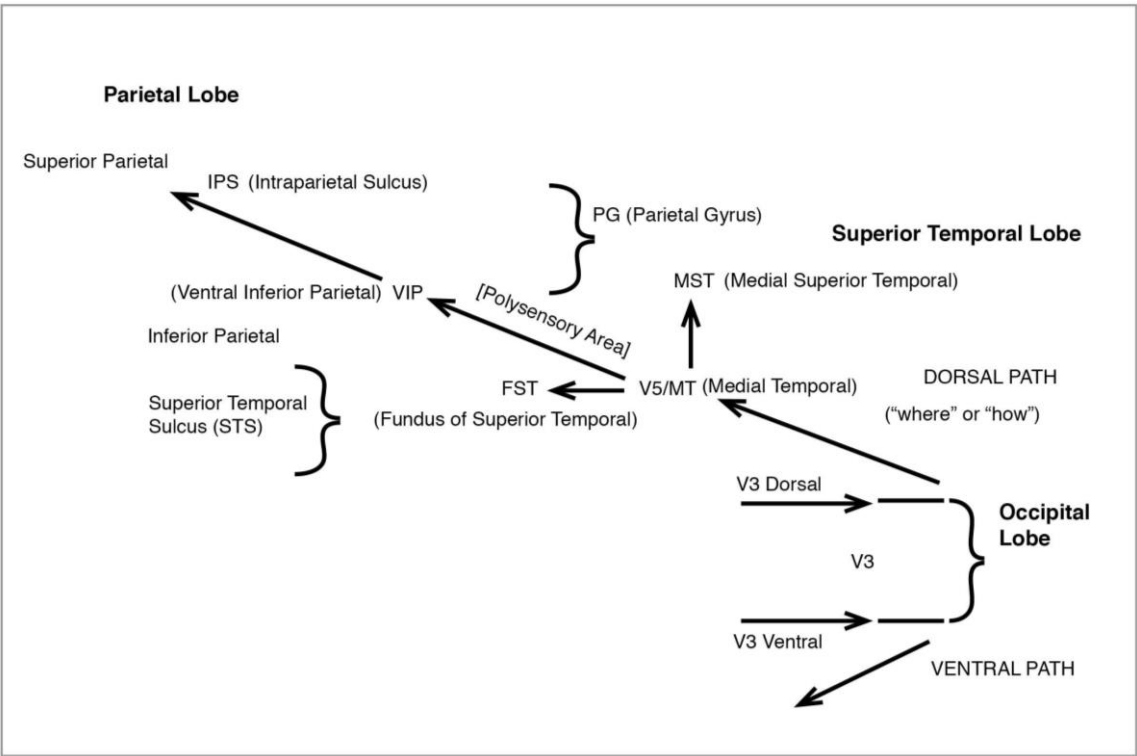


Figure 12. The Visual System: dorsal pathway.



The discovery of two distinct pathways beyond V3 and leading to the high-level areas of the visual system was made by Ungerleider and Mishkin (1982 [34]). The ventral or lower path (Figure 11) proceeds from V3 to V4, then to the inferior temporal lobe (IT), terminating in an area known as TE. This area, like the original parvo, processes colour and sharp foveal vision. It is slower than the other, presumably younger in evolutionary terms, and suited to the precise recognition of objects, not least animate bodies and faces. (For an account of this in connection with rock art, see Dobrez 2010–2011, 2012 [35,36]). The second pathway, the dorsal (Figure 12), like the original magno, processes peripheral vision, is colourblind—and much faster than the ventral. It leads from V3 to V5 (also known as MT), in the superior temporal sulcus (STS), then to the medial superior temporal (MST) and the parietal lobe. I shall have much to say below about the role of MT and MST.

Ungerleider and Mishkin dubbed the two pathways the “what” and the “where”, the one primarily concerned to identify whatever it is we are seeing as precisely as possible (hence “what”), the other to locate visual objects in space (hence “where”). Clearly the latter has more immediate survival value, and this accounts for the greater speed of the dorsal signaling stream. If you are under threat, it is more pressing to register where the threat is coming from than to know exactly what sort of threat it is. That it is a carnivore matters, but not as much as the question “where is it?”. If it is not a carnivore and indeed no threat after all, it has done no harm to press the alarm button. We recall moreover that possible motion in our visual field has pressed a button right from the start, at V1.

The connection between Ungerleider and Mishkin’s two pathways and the earlier magno/parvo streams, as well as the still earlier retinal division between large and small ganglion cells, was proposed by Livingstone and Hubel (1988 [37]) and would seem to have general support (Maunsell and Van Essen 1983, Shipp and Zeki 1989, Boussaoud *et al.* 1990, Goodale 1993 [38–41]). Understandably, however, the original characterization of the two paths has altered somewhat as knowledge has become more detailed. Current emphasis is on form and colour for the ventral as against depth and motion for the dorsal—and it is the issue of motion I now wish to address.

7. MT/V5 and the Processing of Motion

There is wide acceptance that the key area for the processing of visual motion is the middle temporal (MT), otherwise known as V5, in the dorsal pathway to the parietal described above. The term MT comes from study of the New World owl monkey (Allman and Kaas 1971 [42]) which established the zone as visual. Since Old World monkeys are closer to humans, subsequent work largely concentrated on the macaque (and in fact the Ungerleider/Mishkin experiment which discovered the dual-path structure of visual processing used macaques). In the course of such work an MT homologue was found in the macaque (Gattass and Gross 1981 [43]). Zeki (1974 [44]) argued that MT was responsive to movement, and to a greater degree than V1 at the start of the processual chain. He found it directionally selective, sensitive to form in motion and, interestingly, to motion in *any* direction, implying some neurons registering “the presence of movement *per se*” (p. 568). Much of this, in particular the connection with V1, is gone into in further detail in Shipp and Zeki [39], as also in Weller and Kaas (1983 [45]).

However, in 1983 Maunsell and Van Essen published three articles which, with Ungerleider and Desimone (1986 [46]), are especially informative in connection with MT. They argued [38] that in

view of its selectivity for direction and its operational speed, the macaque MT “is more specialized for the analysis of visual motion than has previously been recognized” (p. 1127). They added that this includes the analysis of motion in 3D space ([38], 1983 [47]). The fact that MT had larger receptive fields (that is, access to a larger area of the retinal surface) than V1, meant that it was indeed higher in the processing hierarchy. It could be understood as corresponding, in its dorsal path, to V4 in the ventral (1983 [48]), though Maunsell and Van Essen saw only modest linkage with the ventral form-processing areas. Actually biological motion experiments like Johansson’s on the pickup of form *from motion* suggest the linkage is likely to be more robust—and indeed later neurophysiological work, as we shall see, has posited considerable integration of dorsal and ventral pathways. Nonetheless, in 1983 Maunsell and Van Essen provided a very detailed map of MT’s connections with other visual areas in the brain, in particular its further projections in the dorsal, namely the medial superior temporal (MST) named by the authors and proposed as the next motion-processing stage after MT, as well as the ventral inferior parietal (VIP). We thus have two key processing areas for movement here—MT and MST—leading to VIP which, however, like other upper reaches of the path to the parietal, relates to motion but in a very different way.

Other researchers have underlined the role of MT as probably the main site for motion processing (Albright 1984, Newsome and Paré 1988 [49,50]). The latter, listing its likely functions, conclude that it is a “general purpose motion processor” (p. 2210), adding the clinical observation that MT lesions in fact impair the perception of movement. But it was Ungerleider and Desimone [46] who confirmed the centrality of MT and its connection to MST and VIP and, additionally, to FST, the fundus of the superior temporal. They also pinpointed another stream to the parietal, associated with the parieto-occipital (PO), this last examined by Colby *et al.* (1988 [51]) as making a contribution to the processing of peripheral vision (we recall that the ventral path deals with foveal or central vision). Ungerleider and Desimone’s paper effectively details all the stations for dorsal processing of motion. What begins with sensitivity to movement in V1 continues along the occipital route to V2 and V3, then diverges from the ventral object-processing path to reach several key locations in the superior temporal, to take, finally, a different form in its passage through the parietal lobe. (In fact, as earlier noted by Maunsell and Van Essen [48], the final projection would seem to be to the frontal eye field (FEF), in the direction of the frontal brain.) For Ungerleider and Desimone the MST area, immediately beyond MT, added to the latter by processing more and more complex motion. But there may be more to it than that.

Saito *et al.* (1986 [52]) suggested that the role of MST was to *integrate* motion perception, “to discover whole events of visual motion” (p. 145). This is important, if not unexpected. If the early stages in the visual process analyze basics such as shapes in terms of edge-orientations, that is, places where the eye registers difference or discontinuity (as signaling the border between shape A and shape B), higher levels of analysis must be capable of synthesis. After all, we see movement, say a running figure, “whole”, not as an aggregate of “bits”. We see events, however complex, as *entire* events. Saito *et al.* referred to the MST area, on which they focussed, as DSR because it contained what they termed D, S and R cells, all three classes involving directionally-selective neurons. D cells (D for direction) fired for frontoparallel motion in a straight line, that is, movement across one’s visual field. S cells (S for size) fired for radial motion, that is, movement in depth signaled by size change in the moving object. (I shall return to radial motion later in this article.) R cells (R for rotation) fired for

circular motion. Apparently, D, S and R cells accounted for more than 80% of all cells. From which the conclusion was drawn that MST neurons took up the all-purpose motion information from MT and synthesized specific types of motion. Interestingly, MST included cells referred to as “jerky” because they fired at the start and stop phases of movement, that is, responded both to positive (acceleration or “start”) and negative (deceleration or “stop”). All this indicates fine tuning of motion perception, reminiscent of and parallel to the fine tuning of object perception in the final stages of the other pathway, the ventral. It is matched by receptive field size in MST which is greater than in MT. That means that MST has access to large areas of the retina, something that would be needed for its integrating function.

It is worth adding that the Saito *et al.* article also considers MST as possibly specifically integrating motion with form—not because of a direct connection with the object area in the inferotemporal (which, as explained, permits vision with extreme acuity), but via a “polysensory” link. The polysensory, like MT and MST, is located in the superior temporal STS, but at a stage above these on the processing hierarchy. So we are adding further brain topographical/functional areas which contribute to an integrated visual response to movement in our visual field.

Something like this emerges in Boussaoud *et al.* [40], including all projections from V1, ventral and dorsal, with, in the dorsal path, MT projecting to MST as well as its abovementioned associate FST, then to the various subsequent areas of the parietal and the polysensory area mentioned by Saito *et al.* Actually the Boussaoud *et al.* paper proposes nothing less than a *third* pathway for visuals, though as it emerges the claim is in fact the more modest but still very significant one of MST (or, as it now appears, MST and FST) as integrating motion with form via the polysensory, as Saito *et al.* had thought. At the same time there is tentative talk of a dorsal-ventral link, possibly via TEO in the inferotemporal, the penultimate stage—prior to the terminus TE—for the processing of objects. Any such link between the high-level areas of the dorsal path (MST/FST) and a high-level area of the ventral (TEO) would indeed be significant, as it would begin to give us something like a complete picture of how we see both *motion* and *things*, either in motion or stasis. Since it is probably the case that *both* pathways to an extent process *both* movement *and* objects, even if the dorsal mostly deals with motion and the ventral mostly with objects, it must be the case that, in addition to replicating each other, the two also communicate, and very efficiently. A coordination of form and motion perception as required to explain Johansson’s experiments is of course necessary in some way or other. Boussaoud *et al.* as it happens make specific reference to biological motion in the context of the polysensory, and I intend to return to this below.

It is time to explain what I mean by stating that the dorsal path beyond MST processes motion in a way distinct from the steps preceding it. If MT and MST provide the neural substrates for *seeing movement*, what follows provides the neural basis for *seeing in order to act*. So dorsal functions progressively shift from motion perception to the visuomotor, that is, the provision of visual information enabling performance of motor acts (such as, for example, coordinating eye-movements and movement of arm and hand so as to reach out to pick up an object, or catch a ball, *etc.*). We are now in the area of the inferior parietal (specifically PG, to which MT projects), after which, in the superior parietal, we are, as I understand, as much in the “doing” as the “seeing” area or rather the coordination of the two indicated by Saito *et al.* as “integration... indispensable for the perception of

the relationship between the movements of the self and the visual environment, as well as for visually guided motor control” ([52], p. 155).

Strictly speaking, the uppermost regions of the original Ungerleider and Mishkin “where” pathway are not relevant to the present argument. It suffices to say, for interest’s sake, that they are discussed in Leinonen *et al.* (1979 [53]), Hyvärinen and Shelepin (1979 [54]), Motter and Mountcastle (1981 [55]), Sakata *et al.* (1983 [56]) and Baizer *et al.* (1993 [57]). Issues such as eye-movement relevant to this stage of processing are also investigated, but with reference to earlier MT/MST contribution, by Komatsu and Wurtz (1988 [58]), and Newsome *et al.* (1988 [59]).

As an addendum to this section of the argument, I need to mention Goodale and Milner (1992 [60]) and Goodale [41] which state most clearly, with general reference to primates, that the final goal of the dorsal pathway is indeed visuospatial/visuomotor by renaming Ungerleider and Mishkin’s “where” path the “how”. “How” because the endpoint of dorsal (especially peripheral) vision is not seeing motion so much as seeing motion with the aim of performing an action. In the process these papers question the validity of the magno/parvo, as well as of the dorsal/ventral distinction, or at least complicate it with a reassertion of the strong link between MT (V5) in the dorsal and V4 in the ventral, likewise with stress on ventral input from V4 to the parietal via MT and, finally, with the postulate of a direct connection between the ventral inferotemporal and the posterior parietal:

We assume that the two systems will often be simultaneously activated (with somewhat different visual information), thereby providing visual experience during skilled action. Indeed the two systems appear to engage in direct crosstalk: for example, the posterior parietal and inferotemporal cortex themselves interconnect and both in turn project to areas in the superior temporal sulcus [that is, areas like MT and MST]. There, cells that are highly form selective lie close to others that have motion specificity, thus providing scope for cooperation between the two systems. In addition, there are many polysensory neurons in these areas, so that not only visual but also cross-modal interaction between these networks may be possible. This may provide some of the integration needed for the essential unity and cohesion of most of our perceptual experience and behaviour although overall control of awareness may ultimately be the responsibility of superordinate structures in the frontal cortex (Goodale and Milner [60], p. 24).

It seems to me that, in line with these sorts of comments, time will inevitably reveal more rather than less integration of brain systems, including the visual. Nonetheless what we have in the above quotation already envisaged the business of motion-perception as a complex cooperation between nothing less than all four lobes of the brain, occipital, temporal, parietal and frontal!

At this point we may take it research has established specific areas for the analysis of real motion and mapped its trajectories.

8. Processing Motion in the Human Brain

But all this, except for the discussion in Goodale and Milner, is entirely with reference to experiments on monkeys. What about the human brain? It seems there are basic similarities—as well as differences in the precise location of motion-processing zones. Accordingly, while retaining the terminology developed in the course of monkey studies, researchers speak of “homologous” areas in

the human visual system. Since it is illegal to sacrifice human subjects and slice their brains post-experimentally in the manner done with monkeys, experiments with humans rely on non-invasive PET and fMRI scanning. As I understand it, we still know rather more about the monkey system than the human—so, like others, I illustrate my argument with mapping derived from work on monkeys (Figures 10–12). Still, this is applicable, if in broad-brush terms, and quite adequate for present purposes.

The human visual system has a ventral/dorsal divergence (Tootell *et al.* 1995, Paradis *et al.* 2011 [61,62]), one for colour and form, one for motion. Schenk and McIntosh (2010 [63]), who specifically address this point, underline it, with the proviso that the two paths are not independent of each other—something which comes as no more of a surprise than it did with the Goodale and Milner comment covering primates in general. Testing human reactions to moving as against stationary objects, Watson *et al.* (1993 [64]) found key motion processing (V5 in their terminology) at the occipito-temporal border, with activation extending to the parietal. This would pretty much correspond to MT/MST in the monkey superior temporal, with its destinations in the parietal. Replicating the experiment, Tootell *et al.* [61] agreed that the same area, which they called “human MT”, receives magnocellular input, activates the inferior parietal and is, in short, “functionally indistinguishable from macaque MT” (p. 3218). Work on human brain lesions (Plant and Nakayama 1993 [65]) also points to an MT homologue in connection with the rare phenomenon of “motion blindness” in injured patients who clearly perceive shapes but cannot register movement and so view events as a confusingly unrelated succession of static tableaux. Thus they see a car at point A, then at point B, without registering the movement from A to B. While focussing on illusory motion—itself not relevant to the present argument—Federspiel *et al.* (2006 [66]) come to the same conclusion as regards MT/V5, with the caveat that the precise location of various aspects of motion in the human brain remains controversial. By 2011 general confirmation of this “human MT” or V5 thesis continues (Paradis *et al.* [62]), though it seems we still need to take the research as provisional. As Kaas (2004 [67]) had it, we acknowledge the likely general validity of the thesis with the understanding that the monkey analogy is not unproblematical and that experiments on human subjects have some way to go.

Let us note by way of conclusion that the homology case for MT has been extended to parietal functions in the dorsal pathway. Again on the basis of fMRI scans, Culham and Valyear (2006 [68]) and Farrer *et al.* (2008 [69]) locate the perception/performance, that is, visuospatial/visuomotor areas in the human parietal. Just as motion vision focusses in the monkey MT so as, ultimately, to lead to visually-coordinated action (such as grasping an object), so also in humans, registering motion ultimately leads to those higher parietal regions feeding into motor activity and the awareness of action.

9. Biological Motion and the Human Brain

I have given an outline of neural substrates for visual perception in monkeys and humans. But we should return to a comment (Boussaoud *et al.* [40]) linking monkey cells in the “polysensory” superior temporal area to selectivity for *biological motion*, that is, the movement of living creatures discussed above in the context of Johansson’s form-from-motion point-light experiments. Now biological motion is highly complex motion, and there has been research seeking to locate processing of such motion in a

particular human area. With a nod to Johansson, Bonda *et al.* (1996 [70]) combined PET scanning with point-light displays. Like others just mentioned, they found a motion centre in the occipito-temporal, specifically the inferior temporal sulcus (ITS). In connection with biological motion in particular, they found that goal-directed hand motion activated the superior temporal (STS) en route to the intraparietal sulcus (IPS), which was *especially* activated. This effectively parallels the dorsal route for monkeys: you *see* the motion in the MT zone (within the STS) and especially register response in the parietal, where motion perception, as explained, feeds into visually-directed motor activity. This merely for a hand motion. Expressive whole-body movement registered in the same visual STS/MT zone, with some affective fallout in the form of limbic system activation, notably in the amygdala—the emotive centre of the brain. Obviously the motion pathway needs to be sensitive to the emotional content of perceived bodily motion (anything from desire to terror!).

Grossman *et al.* (2000 [71]) combined fMRI with point-light displays. Using (artificial) coherent motion stimulus, they put the homologue for monkey MT/MST in the human temporo-parieto-occipital junction, in short, an area involving all three lobes with a major stake in the perception of motion. Biological motion specifically registered in a way analogous to the operation of the monkey polysensory—in the human posterior STS (lateral and anterior to human MT/MST). Interestingly the paper makes the incidental observation that monkeys respond appropriately to point-light displays! Another human fMRI experiment in combination with point-light (Grèzes *et al.* 2001 [72]) echoes findings given immediately above, drawing attention to the occipito-temporal, posterior STS and intraparietal sulcus (IPS). Mention should also be made of an experiment with fMRI but using perceived objects in motion and not point-light. Pelphrey *et al.* (2003 [73]) went further than others by comparing (1) a walking human to (2) a walking robot to (3) a disjointed mechanical rearrangement of the robot to (4) a moving, somewhat anthropomorphic grandfather clock. Biological motion in the human and the quasi-biological robot rated better than the rearranged, non-biological robot and the clock. The conclusion of the experiment being that human STS, that superior temporal zone to which the present argument constantly returns, is geared to register biological motion *per se*. In view of which Johnson (2006 [74]) was prompted to refer to the existence of a perceptual *life*-detector!

10. The Neurophysiology of Depicted Motion

So we are geared to see movement—but especially that of living things. It makes perfect evolutionary sense. But where does *depicted* motion fit into the scheme? Since we register depictions of motion, we must have neurons for the job. Earlier in this paper I began with psychology experiments bearing on biological motion perception, then proceeded to similar experiments dealing with “implicit motion” and Representational Momentum—in the present context: depicted motion. I now intend to duplicate that structure in the context of neurophysiology, progressing from neural substrates for biological motion perception (just considered) to substrates for depicted motion. As it happens, these last have also been identified.

Senior *et al.* (2000 [75]) used fMRI to locate a brain region for Freyd’s RM. Comparison of firing for a depiction of objects at rest *vs.* video of objects in motion pointed to V5 as the motion area—nothing new here, given my preceding comments. But comparison of firing for pictures of objects at rest *vs.* stills *prompting RM* also pointed for the most part to V5. Thus the same area which

responds to real motion responds to Freyd's "implicit" (read "depicted") motion. The experiment suggested high-level processing for both real and depicted movement along the dorsal path to the parietal. Senior *et al.* saw a need for semantic input into the process, with some contribution from the ventral object-analysis path which would supply the long-term memory involvement they took to be required—required to know, in the light of the Reed and Vinson experiment, that rockets move and steeples do not.

At the same time as the Senior *et al.* paper, Kourtzi and Kanwisher (2000 [76]) did a very similar study which activated the MT/MST (that is, V5) area, showing real and depicted motion registering in the same region. They too thought that such high-level analysis as that required for a *picture* of motion could not be limited to this area. Again, since object recognition was involved, they postulated long-term memory assistance from outside the motion processing zone. The fact that showing pictures of houses (generally at rest!) evoked no response from observers, suggested cognitive inference: we *know* that houses, like steeples, stay put. And yet the visual processes in question were quite automatic.

Overall, differences between the two experiments do not appear significant, though mutual critique ensued (David and Senior 2000, Kourtzi and Kanwisher 2000 [77,78]). What is striking is that both lots of experimenters wanted to query the capacity of the motion path to process stills as moving, while noting evidence that it did just that. David and Senior speculated: "we are forced to the rather counter-intuitive conclusion that... the V5 complex may in fact be smarter than we think and that experience of the world of movement as well as semantic knowledge may be encoded and stored within the visual system itself" (p. 1365). This would not exactly contradict the Reed and Vinson line, but it would tend towards Freyd's proposal of the visual as sufficient unto itself. Something of a similar ambivalence was evidenced by the Senior and Foley online report [27] mentioned earlier in this paper. Before seeking some resolution to the issue, however, I need to add an important piece of information. As the 2006 report notes, those 2000 experiments by Senior *et al.* and by Kourtzi and Kanwisher showed that V5 was centrally involved in depicted motion processing, but not that it was *essential* for that processing. Senior *et al.* (2002 [79]) set out to demonstrate its necessity, and did so by using transcranial magnetic stimulation (TMS) to complement the usual fMRI scan. TMS created a magnetic field, in this case in V5, which disrupted neural firing, that is, simulated the effect of a lesion in V5. It amounted, in fact, to a—harmless and repeatable—"virtual lesion" (p. 84). The result showed that V5 disruption inhibited the RM response, thus proving that V5 is necessary for the generation of RM. In terms of the present argument this means that tampering with the movement area in the dorsal path *prevents* our seeing a picture of a running figure as running.

11. Percept and Concept Revisited

The cognitive question has of course been raised earlier in this article, but I want to return to it, this time with an eye to neurophysiology. Senior *et al.* [79] simultaneously postulate need for contributions extraneous to V5 (such as frontal cortical regions) for the processing of depicted motion—and processing *within* V5 itself: "an alternative (and more radical) hypothesis could be that the V5/MT system is able to carry out 'high-level' processing, necessary for RM, on the basis of previously 'learnt' amodal information (semantics)... this process... carried out without the need for frontal or 'central executive' resources" (pp. 90–91). Processing within V5 on the basis of learned information is

perfectly possible, since individually learned facts are held in long-term memory in the same areas used for their perception (see Dobrez and Dobrez [10] and 2013 [80] on the part played by such memory in the recognition of objects). But this would hardly apply to the experience of gravity. It would be very odd to suggest that each time we see Freyd's subject jumping from his wall we access a long-term memory of gravity—in order to *see* the jump, that is, register it as depicted motion. A more likely perceptual scenario is that knowledge of gravity is (following Shepard's argument) hardwired and that we assume this to be the case when we judge the perception to be "automatic". That hardwiring could, of course, be as Senior *et al.* envisage, in V5. Putting it another way: perception is its own kind of cognition, or, still more pithily, perceptions are intelligent. This thesis is very much in line with Zeki voicing suspicion of our everyday separation of "seeing" and "understanding" in his book *Inner Vision* (1999 [81]). After all, we say we "see" the object in our field of vision—and when we get the point we also say we "see" it. So Occam's Razor would suggest that Freyd's case for the perceptual nature of the RM phenomenon holds. In the context of neurophysiology, the evidence is that V5 knows what it is doing!

In the context of psychology above, I attempted to negotiate the diverse conclusions of Freyd and the Reed and Vinson experiment at least in part by appealing to a perceptual attentional "switch". Here I refer the reader to O'Craven *et al.* (1997 [82]), who found that voluntary attention shifts modulated firing of neurons in the human MT/MST area. The experiment was simply with moving and stationary dots, with attention shifting from one to the other and firing rates increasing with attention directed at moving dots. It is not an experiment directly relevant to the present argument except for the fact that it suggests a parallel: if attention can shift from static to moving objects and back again, then it can equally register movement in a picture and freeze it at will—this being the case with Hockney's leopard. But there is a more important point. After all, we need no experiment to tell us that we can shift attention from something moving to something static, *etc.*—or from the same (depicted) object *as* moving to it *as* static. The interest of the O'Craven *et al.* paper is that its focus is neurophysiological—MT/MST in particular—in relation to voluntary attentional modulation. By switching attention we increase or decrease neuronal firing rates. It is probably along these lines that a case might be put for semantic intrusion, in the form of a verbal cue ("it's a rocket"... "it's a steeple"), into the operation of MT(V5)/MST. But it would be a case making allowance for the fact that, in the words of David and Senior, V5 might be smarter than we think, that is, that more high-level processing, adequate to such complexity as the perception of depicted movement, is done in V5 than we realize.

12. Revisiting V5

Having said which, it may also be as well to acknowledge further nuance in the discussion. Logic works by the principle of non-contradiction, an either/or, but the brain, one suspects, is more liable to both/and operations.

When Jarrett *et al.* (2002 [83]) found RM heightened in schizophrenic subjects, that is, depicted motion registering in more pronounced fashion than in normal subjects, they interpreted it as a failure of inhibition, an over-the-top automatic response—V5, we might say, let loose in the absence of frontal brain control. The implication being that frontal areas play a part in RM operations, a fact

confirmed by Rao *et al.* (2004 [84]). So while V5 can and probably does do the essentials of the job, it does not work alone.

The most revealing experiment in this connection may be that of Krekelberg *et al.* (2005 [85]) which, as it happens, does not implicate the frontal lobe but just about everything else. Researchers used an fMRI adaptation which offers greater than usual possibilities, though with possible risks. Subjects were exposed to so-called Glass patterns with both radial and concentric implicit-motion (referred to as “implied motion”) effects, and the result indicated some activity in the early visual system and significant response in the dorsal MT(V5), as expected. It also indicated activation in the occipital-temporal: an occipital area termed LOC, plus V3v (“v” for ventral, that is, lower-pathway V3), specifically an area termed VP. And last but not least, V4, that key stage in the ventral path corresponding to and connecting with V5 in the dorsal.

Now the incredible cross-wiring of the brain cannot surprise. I noted above, in connection with real motion perception in monkeys, that it seemed that all four cortical lobes combined to produce that effect. It now looks that, for all the undoubted centrality of the V5 dorsal motion area, all four lobes of the human brain, the occipital at the back, the temporal on the side, the parietal at the top, and the frontal at the front, jostle for a part in the drama of perceiving a static shot of a man jumping from a wall!

The Krekelberg *et al.* reference to the early visual areas brings to mind a comment by Zeki [81] made in the context of Op Art. An Op Art abstraction gives an illusion of movement, shifting and shimmering as we observe. Now illusory motion is not depicted motion. Though there is none in the Op Art picture, we *actually* see motion. Nonetheless Zeki’s comment might be very much to the present point. He states that when we observe *real* movement we register in V1 and V5, that is, at both the initial and the high-level areas. But when we have the *illusion* of movement, as with Op Art, the effect is largely in V5. This suggests that if we largely bypass V1 when observing a picture which generates *illusion* of movement, we may also be liable to do so when observing a picture which simply *depicts* something moving. It also suggests a possible definition of depicted motion as perception which activates V5—without too much input from V1. Of course to the extent that V1 is the gateway to the entire visual system and responsive to moving things, one would expect its involvement. It is just that with depicted motion that involvement might be significantly less. I put this forward tentatively on the basis of Zeki’s expertise and as a plausible way of approaching the whole neurophysiology of depicted motion.

However interesting this may be, though, it is not at the heart of the findings of Krekelberg *et al.* What really emerges from these is the sheer detail of both dorsal and ventral input for implied motion (in the present context, read “for RM operations”). In the dorsal V5, 45% of cells are selective for *both* actual *and* implied motion. Now, as we have seen, earlier studies showed that real and depicted motion are indeed processed in the same MT(V5) *zone*. What this study shows is that the same *cell* in MT(V5) responds to both real and implied motion! Extrapolating from this and other experiments discussed above to the case of depictions, we might postulate that the dorsal makes no distinction between actual and pictured movement. On the other hand, the ventral, V4 in particular, while also registering both actual and pictured movement, does not do so in the one neuron. Thus the ventral makes a distinction between actual and pictured movement. Krekelberg *et al.* posit interaction between dorsal and ventral to account for perception of their “implied motion” and it may be this, along with the Zeki-prompted

scenario just outlined, which offers the most nuanced and probable account of the neurophysiology of perceiving movement in a still picture.

13. The Phenomenon of Pictures

The above, along with belated consideration for the patient reader, encourages me to the following summary of the complicated process I have been investigating. When V5 is faced with real movement it is, if we may so put it, happy to fire. If it is faced with a *picture* of movement, it fires *anyway*. What Shepard [14] referred to as a “resonant system” is able to be excited in different ways, including “by a signal that is slightly different, weaker or incomplete” (p. 433). Perhaps this is a case of what I termed the brain’s near-enough principle: if it looks somewhat like X, we treat it as X; if it looks near enough to the real thing, we treat it as the real thing. After all, there is no evolutionary risk in this; it is the safe option. Now when the ventral path is presented with *real* motion it extracts its form (the form of what is moving). So it underpins Johansson’s form-from-motion phenomenon, which we accordingly source to the ventral pathway. When the ventral is presented with *depicted* motion it zeroes in predominantly on form, just as analogously the dorsal zeroes in predominantly on motion. But it is the dorsal which may be presumed to register *motion-from-form* in the case of a picture, that is, read Hockney’s leopard as moving—a highly sophisticated operation; while the ventral registers *form-from-motion*, not in a picture but in the case of actual movement. With a picture, the ventral presumably analyses the form with greater precision than the dorsal is equipped to do.

I want to add here that such considerations, especially in the light of the near-enough/good-enough principle, raise nothing less than the issue of pictorial representation. The Saito *et al.* study uncovered neurons that responded only to real moving objects and not at all to projected ones. So, unsurprisingly, we can differentiate between the real thing and the picture. But there is considerable evidence, from neural studies (above), and from observation, that the two are able to be run together. This must be the case when we see movement in a picture which is not actually moving. If the dorsal lumps together real and depicted—depicted taken as near enough to real—then depiction *itself* is made possible, if not necessary, by the near-enough principle. A picture of X may be taken as a stand-in for, or an equivalent of, X. This would provide the neural basis for “likeness”, or verisimilitude in art, that is, a picture of a horse that looks like a horse or a portrait that looks like its subject—as well as for depicted movement that looks like actual movement. *The near-enough operation of the visual system allows for the phenomenon of pictures.* All this still leaves open the question of why we depict, that is, make pictures, and the answer is probably an evolutionary imperative of communication within the human group.

Be the above as it may, investigation into the neurophysiology of motion has brought us to a provisional but likely definition of depicted motion in neural terms or, putting it another way, to a likely substrate for the strange but everyday phenomenon that we see a galloping horse in a picture *as galloping*. Somewhat unexpectedly, it has brought us to a neurally-based and not implausible preliminary definition of depiction itself.

14. Event-Perception: “Arm’s Length” and Looming

So far we have been discussing perception of motion with motion generally assumed to be of the frontoparallel sort, that is, as passing *across* the observer’s visual field. It is, however, time to divide motion perception into two major categories of special importance both in life and pictures—and this for reasons of evolution. The first is motion perception which does *not* involve the observer, the second motion perception which very much involves the observer. If I watch something going on without engagement on my part, perception entails no risk. I may be looking at a herd of animals, keeping track of their movement, either for hunting purposes or simply to avoid them. The moment one of them turns in my direction and approaches, with or without aggressive intent, perception has action consequences and entails some kind of risk: at the very least I am required to make some decisions. At this point the visuomotor/motor systems come into play. Clearly both perceptual situations *matter* in evolutionary terms. But the second may be expected to take priority, since it matters specifically to *me*, the observer. I refer to these two instances of motion perception as “arm’s length event perception” (viewing movement *across*) and “looming event perception” (viewing movement *towards*). These may be related to two recurring types of depictions and it is in this context that I want to consider them. The situation being that while the first is discussed in Art History and rock art studies, but without serious analysis, the second has never been isolated, given some definition or even named. I shall proceed by outlining the phenomenon of visual “looming”.

Looming features perceptually with “optic flow”, that is, the advancing of objects towards me, to pass by on either side as I move through a landscape. However, my interest here is a variant of this phenomenon, namely an object coming towards me, that is, getting bigger and bigger as it takes up more and more of my visual field until, theoretically, a collision occurs. Looming is this phenomenon of impending collision. Before examining what form it might take in pictures, I need to consider it in the context of perceptual psychology and neurophysiology.

15. The Phenomenon of Looming

Gibson [4] gives a clear description of the phenomenon. Naturally, as with other researchers, his interest is equally in the opposite of looming, namely an object receding from view. Thus it is a case of magnification/minification as the object takes up more and more of the visual field or less and less of it. My concern is solely with magnification, and this because of its pressing evolutionary implications. Visual recession doubtless matters when you are chasing something, but its opposite will be a matter of life and death when something is chasing *you*. Gibson sources the term “looming” to Schiff, and gives an account of the experiment done by himself with Schiff and Caviness (Schiff *et al.* 1962 [86]) on its effect on monkeys. Unsurprisingly, looming evokes alarm. Moreover such alarm is innate or hardwired, and in humans too, since, for example, knowledge that no collision will ensue has no effect on the alarm response: watching a 3D movie you duck, mentally or literally, though you know it is only a movie. Gibson’s 1979 account is of some interest:

Experiments showed that the size and the distance of the virtual object [a progressively magnifying shadow on a screen] were indefinite but that its approach was perfectly definite. After the shadow filled the screen, the virtual object seemed to be “here”, at zero distance. It did

not look like a shadow on a screen but looked like an object. The object in fact came out of the screen. This was only to be expected, for, by the laws of natural perspective, the closer an object comes to the point of observation, the closer its solid angle will come to a hemisphere of the ambient array (that is, light reaching the retina).

There seemed to be a direct perception of an event that could be described as *approach-of-something* ([4], p. 175).

The literature of perception makes frequent allusion to looming (Shepard [14], Neisser [17], Grossman *et al.* [71]). It has been experimentally linked to RM by Kelly and Freyd (1987 [87]). The authors begin by distinguishing between, on the one hand, simpler motion such as an object's rotation or its translation (that is, frontoparallel motion) and, on the other, more complex motion in depth. Now we will perceive *some* motion in depth as part of an "arm's length event", that is, figures moving, to an extent, in depth as they pass across our field of vision. But I shall reserve the terminology of motion in depth for the looming phenomenon which relates *entirely* to depth. Echoing Gibson, Kelly and Freyd place looming in this context, as a "transformation" or change of an object in the third dimension: "proximally, this transformation is specified by a change in the projected size of an object. However, distally this transformation represents an approach or recession of an object". Since the same laws must apply here as with other forms of motion, "one might expect momentum effects in this type of situation if the visual system responds to the size changes as indicating motion in depth" (p. 384). In short, if we expose the observer to a magnification effect this will be read as, in Gibson's words, "*approach-of-something*". Increase in the size of an object will not cue "something getting bigger" but, for excellent evolutionary reasons, "something getting *closer*".

In which case RM should come into play, as it did with other forms of movement. Kelly and Freyd experimented with "grow" and "shrink" squares and the former—the relevant ones for me—were indeed remembered as larger than they were. That is, perception overshot the mark, as it had for the stills of Freyd and Finke's "rotating" rectangles [18]. Notably none of the 16 experimental subjects reported squares progressively increasing in size. Rather all saw an approaching object. Other experiments were undertaken to rule out the possibility of the results being attributable to size *per se* and, with modest proviso, the authors concluded that the looming effect did in fact involve RM.

Now for Kelly and Freyd a *momentum* finding mattered, since that gave support to the thesis that the motion in question related (however "abstractly") to an internalized Newtonian law. For the present thesis, focussed as it is on depicted motion—at this point, of the looming variety—it makes no difference whether the displacement revealed by Kelly and Freyd is specifically one of momentum; a finding of visual displacement suffices for my purpose. So the fact that the experimenters were not totally certain that *size* of itself might not trigger the looming effect is allowable as far as I am concerned. My interest is in whether a Freyd-type experiment shows that a looming image in a still picture does indeed register as moving. Having said which, it is as well to return to the conclusion drawn by the experimenters: that it *was* a case of RM.

However interpreted, the Freyd experiments lead to an important, though expected, conclusion for the present argument. What was said above concerning perceived motion in static compositions applies to the situation of looming. In other words looming, read as movement towards the observer, measurably obtains when the latter is viewing a static picture of a looming object. I shall analyze

this type of depiction further, but first want to suggest likely neural substrates for *real* looming motion.

16. Neural Substrates for Looming

As recently as 2013 Tyll *et al.* [88] admit we know a lot more about the monkey than the human brain. In the same way as I did above when discussing neurophysiology I shall rely heavily on experiments with monkeys, noting, however, that researchers of the looming phenomenon have a predilection for the locust, amply studied for its loom-sensitive neurons, the lobula giant-movement detector (LGMD) and descending contralateral-movement detector (DCMD)! The reader is referred to Rind and Simmons (1999 [89]), which compares collision-prediction for the locust and the pigeon. With regard to monkeys, it seems that looming has chiefly been traced not to the MT motion area but to its subsequent MST stage, somewhat higher on the dorsal visual ladder. This makes sense in that looming prompts an action response: “duck!” or “get out of the way!” and so may be expected to register at higher levels in the direction of the visuomotor parietal. Accordingly Maunsell and Van Essen [47] found that although MT cells process 3D information, they are really geared to frontoparallel and not movement toward (or away from) an animal. Similar results came from Ungerleider and Desimone [46]: MT works on information from the tangent plane, while MST is specialized for depth motion. Saito *et al.* [52], we recall, found MST cells (with large receptive fields) which were directionally selective. There were D cells (54.4%) for frontoparallel movement, S cells (15.7%) for radial movement (that is, in/out) and R cells (13.7%) for rotation. Of course present interest is in the S cells, since “movement-in” constitutes looming. Relevantly, the researchers conjectured that while S cells could be size-change detectors, their key role was detection of motion in depth. Again, we recall from psychology experiments discussed above that size change in the way of magnification is inevitably read as motion in depth, in short as an object looming at the observer.

The above conclusions are supported by Boussaoud *et al.* [40]. Beyond this, Duffy and Wurtz (1991, I [90]) supply additional detail, concentrating on a dorsomedial portion of MST (MSTd) and in particular bolstering the Saito *et al.* argument, though with different figures. They distinguish single-, double-, and triple-component neurons, that is, neurons selective for one, but mostly two or three components of optic flow, namely planar, circular and radial motion. This means that most MSTd cells respond to two or three types of movement, as for example planar (frontoparallel) and circular (rotation)—or circular and radial (in/out), *etc.*—or all three. In the experiment 15% gave exclusively radial responses, which of course includes looming phenomena. However, the thrust of the Duffy and Wurtz argument was in the direction of neuronal combinations primarily aimed at the processing of optic flow, namely the movement of things we register as we move through a landscape—rather than the “approach of something” which is of interest here. Other studies (Sakata *et al.* [56] with reference to the parietal area PG, Anderson and Siegel 1999 [91] with reference to the aforementioned polysensory area) implicate areas still higher than MST in the looming effect. But these focus on preliminaries for action of one sort or another, the visual tracking of a moving object (smooth-pursuit eye movement), the coordination of vision and our own motion as we traverse a landscape. Even in these visually-high neural latitudes, though, perception of motion in depth prioritizes *approaching* over receding motion (Sakata *et al.* [56]). I repeat the warning that all of these experiments are on monkeys

and not humans. Still, we may expect the usual homologies. Curiously, in this connection, Maier *et al.* (2004 [92]), who investigate looming effects on monkeys, report an equivalent of Freyd's RM in the *human* response to "looming" *sound*. Just as we overestimate *visual* perception of looming motion, in order to minimize risk, so we overestimate in predicting an approach signaled by sound. Auditory perception interprets increasingly *loud* car noise as closer (to contact!) than it actually is. In fact Kelly and Freyd [87] foresaw just such "auditory momentum".

17. Event Perception and Narrative Depiction

So much for the phenomenon of looming in psychology experiments and in neurophysiology. With regard to that other event perception which I termed "arm's length", a psychology- and neurophysiology-oriented account has of course already been given, since, as pointed out, "arm's length event" perception involves viewing movement *across* one's visual field, as distinct from looming as a coming-*towards* the observer. The subject of depicted motion has been introduced above for *both* types of movement, showing, in psychology experiments, the operation of the "implicit motion"/RM effect in each case and, in neurophysiological work, probable neural locations for it. In what follows I shall focus on perceptual analysis, along phenomenologically-descriptive lines and, to an extent, formal analysis, of the two types of depicted motion or depicted event. My point is that since "arm's length" and looming event perceptions are situationally so important as to have been hardwired (as evidenced by neural structures just discussed), we may take it that there are likely to be pictorial categories corresponding to them. These will be ways of depicting which are universal, that is, which occur non-culture-specifically over time and place precisely because they relate to and emerge from just those hardwired perceptual situations. We may expect to find these pictorial types in any form of art, ancient or modern. However, given the time depth of rock art, it is reasonable to give greater attention to it in the ensuing discussion.

Now a universalist account will of necessity concentrate on certain general features, discounting particularities explicable in historical terms. Furthermore, in concentrating on depiction of motion, I do not mean to suggest that other perceptual/pictorial categories do not exist. In fact others may be identified, notably the "canonical" (Dobrez and Dobrez 2013 and forthcoming [10,80,93]) which concerns the recognition of objects. This is not of present concern, though it is worth observing in passing that recognition of things by their "canonical form" has its counterpart in the recognition of things by the way they move, that is, their "canonical motion". In the past I described the "arm's length event"/looming event or no-risk/risk pictorial categories as Dynamic and Hieratic, or Narrative and Hieratic or, more recently, Narrative and Performative (Dobrez 2007 [94], 2008 [95], 2010 [1,35,96], 2011 [36,97], 2010/2011 [98]; Keyser *et al.* 2013 [99] and forthcoming [100]). Here I shall use this last terminology and begin with an account of Narrative Depictions which combines (somewhat modified) approaches previously used with the experimental material considered in this article.

I have previously defined event perception in life and art as perception of "something going on" or "something happening". With these guanaco from Charcamata, Argentina (Figure 13), there is no indication of activity. On the other hand, the guanaco from Cueva de las Manos, Argentina (Figure 14), are clearly doing something, say running or leaping. What precisely they are doing is not relevant here, since concern is with the *fact* that we register an event, not with the *content* of the event.

There will of course be borderline representations and this is doubtless the case with the example of a significant association between two bison facing each other at Niaux, cited at the beginning of this paper (Figure1). Do we recognize their “facing” as purely formal, a matter of compositional arrangement, or as “something happening”? I would be inclined to say that this example falls short of a depicted event. The example of the macropod and dogs association from Yankee Hat, also cited above (Figure 2), is also borderline but perhaps just makes the grade. These could, at any rate, be the subject of a cognitive experiment.

Figure 13. Charcamata, Santa Cruz, Patagonia, Argentina.



Figure 14. Cueva de las Manos, Santa Cruz, Patagonia, Argentina.



Just how the “something happening” definition dovetails with scientific evidence should now be evident. Perceptual studies show what we might have expected, that we do not see “X doing something” so much as the *doing*, the action itself; not simply the figures in an event but the event itself—and this in art as in life. We recall Freyd’s work which for my purpose applies Johansson’s conclusions to *depiction*, her insistence on “perceiving transitions”, that is, temporal action: “we pick up real-time change and... parse the world into events that extend over time” ([19] p. 427). We also recall comparable assertions by Schütz-Bosbach and Prinz [20], all this being connected with our capacity to *anticipate*, that is, predict future states of affairs. When RM operates and we read “something happening” in a picture we do exactly that: we anticipate. Now an event is a visual synthesis. It requires, as Saito *et al.* [52] argue, neural integration, the capacity to see “whole”. When we observe the active guanaco from Cueva de las Manos we see action whole, not as an aggregate of bits that go to the making of an action. Putting it another way: we *see* the event; we do not deduce or infer it. This fact is implicit in those point-light experiments and explicit in Gibson [4]. The brain has no time for laboured operations and neither has the survival imperative. Rather, neural mechanisms cut corners and feed-forward information asap. In short, we have become hardwired to pick up an event as a complete perceptual unit. We see events because what goes on around us *matters*—and this carries over into readings of pictures. What is “going on” may be an issue of fight or flight, actual or at one remove in communal exchanges. After all, we are extremely social animals and rely on communication: to read a state of affairs, that is, an event, is vital for us. As argued above, the fact of communality may well be at the root of depiction.

18. The Interaction Fallacy

Once we absorb the straightforward notion that a depicted event is a perceptual “happening”, we are in a position to question problematical assumptions current in rock art research. I shall term the first of these the Interaction Fallacy.

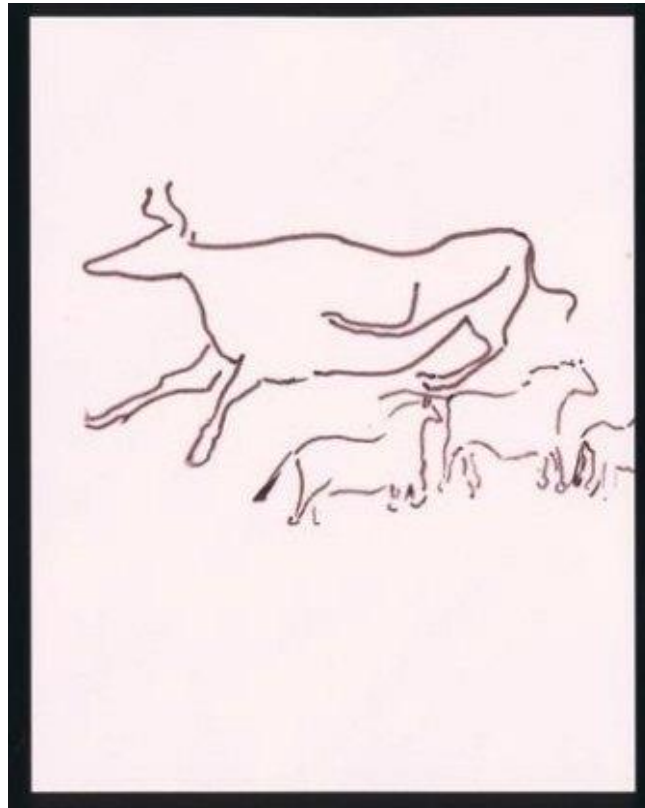
The Interaction Fallacy, common in rock art research, states that a depicted event (a “scene”) requires interaction. This is not necessarily so. It further states that interaction requires the involvement of two or more figures. Again, this is not necessarily so. To readers unfamiliar with assumptions made in rock art studies, what follows may seem a stating of the obvious. Depiction of an event may be of

- (a) a single figure doing something (an action: for example “figure running”)
- (b) a single figure having something done to it by a non-depicted agent (an implicit interaction: for example “figure hit by spear”)
- (c) two or more figures doing something with or to each other (an explicit interaction, either as figures doing something *together* or *oppositionally*).

The so-called “leaping cow” in the Lascaux Axial Diverticule (Figure 15) is, as far as one can tell, doing something on its own. The so-called “wounded man” examples from Cougnac and Pech Merle in the Lot region of France (anthropomorphs apparently pinned with spears) are at the receiving end of an activity. The lined guanaco from Cueva de las Manos (Figure 14) are running or leaping *together* to escape tiny (*oppositional*) hunters on the right. All three pictures are of events or scenes: a cow

“leaping”, men being “wounded”, guanaco “hunted” by men. The first depicts an action, the second an implicit interaction, the third an explicit interaction.

Figure 15. “Leaping Cow”, Axial Diverticule, Lascaux, France.



To return to experimental science: Freyd’s initial example of an event was a single figure jumping from a wall. Dittrich’s [6] point-light experiments included (a) single-figure “locomotory” (walking, jumping, *etc.*) and “instrumental” (hammering, lifting, *etc.*) (b) two-figure “social” (dancing, greeting, threatening, *etc.*).

What is the significance of this? Only that rock art studies assume the Interaction Fallacy stated above, that a scene must contain two or more interacting figures. Lenssen-Erz (1992 [101]), in an article with which I largely disagree, but respectfully, since this is the only work I know on rock art which sets out seriously to ask what it is that constitutes a pictured scene, accepts the usual definition. He mentions rock art scholars who promote the definition and indeed might have added many names to his list. In the course of this he also quotes some notable dissent, namely a thesaurus and a dictionary. The first sensibly gives a scene as “l’être humain... *seul ou non*... représenté en action” (p. 89, my italics); the second, as “any incident or episode” (p. 89). This is more or less what one will find in any dictionary, where the reference is generally to theatre or film—in which Hamlet’s monologue, 007 ducking a bullet from an unseen gun or driving a sports car at speed are taken to be scenes. I have argued against the Interaction Fallacy in the past and if I labour the point here it is because the fallacy is deep-rooted in rock art work. In fact its definition of a scene is arbitrary. More particularly, the definition has no perceptual basis, as evidenced either by visual analysis or experiment. It seems preferable to take a different view and say that a single warrior detached from his frieze at Ubirr, Arnhem Land (Figure 16), and the “wounded man” from Cougnac or Pech Merle come

as much in the category of depicted event as the complexly grouped warriors at Tandjiesberg, Orange Free State (Figure 17).

Figure 16. Frieze of warriors, Ubirr, Arnhem Land, Australia.



Figure 17. Tandjiesberg, Orange Free State, South Africa.



19. Scene as Intended and as Received

At the start of this article I noted the likely stability of the visual system over long periods of time. It is now necessary to raise the issue once more and to consider possible complications. I wish to do this in the context of a continuing analysis of the nature of a scene. As pointed out above, we *see* an event, or, in rock art terminology, a scene. At the same time it has to be added with a different emphasis that it is *we* who see the scene. Of course those who see the picture are not only the ones who *make* it, and any consideration of depictions requires our taking into account both poles of the depictive process, namely the making and the viewing. In the language of hermeneutics these are the poles of original intention or intent (what the makers had in mind in connection with the picture) and reception (what a variety of subsequent viewers had or have in mind in connection with the picture).

Nineteenth-century hermeneutics, beginning with Schleiermacher (1977 [102]), systematized interpretation theory by appeal to the author of the text—which might be verbal, visual, *etc.*—as validating the meaning of the text. To understand the text is to understand the author’s intention in its place and time—a premise which has subsequently constituted the basis of all historical studies, including historical studies of rock art. Twentieth-century hermeneutics problematized the criterion of intent, sometimes constructively, sometimes not. French post-structuralism led down the Derridan path of “undecidability” which simply relativized interpretation, in some instances turning the hermeneutic act into what Barthes called *jeu* or play. Interpretation was entirely a matter of reception, which at its logical extreme might amount to the text meaning what any historical reader or viewer took it to mean. Now the fact of historical reception cannot simply be dismissed. I can be sure that, however much I study the English Renaissance, I will never understand Shakespeare’s plays in quite the way either its author or its audience did. I can be still more sure that I cannot understand the images at Chauvet as did its original makers and original receivers. Does this mean that only original understanding qualifies as understanding? Or, conversely, that since most receivers are removed from the historical situation of the makers, all subsequent interpretations of the text are equally valid?

Observation suggests that both of these options miss the nature of the process in question. In the hands of a scholar like Gadamer (1993 [103]) reception is given due weight in an impressively constructive way, the argument being that, while retaining the principle of historical reconstruction, we also analyze the basic principles by which such reconstruction of the past is necessarily carried out according to criteria of contemporary relevance, that is, viewed through contemporary eyes. Moreover the standpoint of contemporaneity is not something to regret or to seek to neutralize. Rather it is a prerequisite for proper understanding, understanding being possible only when we make past events *present*. If Shakespeare’s plays or the images at Chauvet have no relevance for me, here and now, understanding is reduced to antiquarian curiosity. Such curiosity observes its object from the outside, as an exhibit; it resists grasping the object in such a way as to make it meaningful. For Gadamer, comprehension of some kind may obtain whenever a connection, tenuous or otherwise, exists between two historically separate “horizons”. And there is always the possibility of this, because human experience is not sealed in a series of historical compartments. Or, putting it another way, there are no strictly separate interpretative horizons, only a single, “moving” horizon.

Gadamer’s phenomenological analysis of how reconstruction *actually* operates gives thoughtful weight to the, not necessarily rival, claims of both original intent and subsequent reception. As a philosophical argument, it does not deal with the practicalities of the hermeneutic situation, but any contemporary anthropologist who has real (and not merely professionally-convenient) respect for what she is observing, while being at the same time reflexively critical of her own position, will be putting something like a Gadamerian principle into practice. However, this is not to say that rock art researchers are consciously aware of the legitimate hermeneutic role of reception. They necessarily rely on reception, that is, their present horizon of understanding (they cannot avoid doing so), but since they are doing it unawares they are liable to misuse reception. For example, they may read ancient images in terms of their own implicit, not necessarily critical, knowledge of recent art, and this long after those celebrated blunders made by rock art pioneers like Breuil (who, for example, read “twisted perspective” as a failure of realism, something its makers could not possibly have done). Reception is present as a bias even as researchers downgrade the idea of reception as merely subjective and

overstate our modern incapacity to interpret past images. Much of this may be sourced to the fact that rock art researchers generally adopt a too-straightforward appeal to original intention as the criterion for correct interpretation, that is, follow nineteenth-century hermeneutics. Keenly aware of the practical difficulties of historical reconstruction, they are largely unfamiliar with the way twentieth-century hermeneutics has pointed up its theoretical difficulties. Thus they continue to think of reception as getting in the way of understanding, rather than as assisting it, and consequently fail to take it properly into account—which results both in that unconscious (and therefore unhelpful) appeal to reception and the negative attitude adopted towards it. Following Gadamer and turning from the verbal example of a Shakespeare text to the visual, we might say that it is perfectly reasonable to seek to ascertain what was in Michelangelo's mind when he painted his Last Judgement behind the altar of the Sistine Chapel. We can know something (not everything) of this because there is available ethnography. On the negative side, we can only know *something* and, even then, we cannot know with *certainty*. On the positive, Gadamerian, side, we can know something not merely because we have information, but because we share, to a greater or lesser extent, Michelangelo's historical horizon. But where does this leave the possibility of a contemporary understanding of Chauvet? In the case of rock art ethnography may or may not be available, and usually it is not. Do we have enough of a common history, even with ancient hunter-gatherers, to supply at least a little of the deficit?

Different researchers of rock art will develop different strategies to cope with the problematics of knowing the past, and I do not intend to go into these here. I simply make an observation which is immediately relevant to my argument, namely that it follows from a Gadamerian understanding of historical horizons that the further we are removed in time from a given body of rock art, the more difficult our reconstruction of the mindset of its makers and the more imperative that we take reception into account. Indeed there may be ways of entirely sidestepping the criterion of original intent and turning reliance on reception into an unequivocal positive. This may be done by reading rock art only in the most general terms, as for example by reference to perception—the method adopted in the present article. However, my suggestion is not a prescriptive one. There are, as stated above, varied strategies for dealing with rock art, indeed with all depictions and the entire body of past human experience. These must include the option, where offered, of historical reconstruction.

To return to the matter of a visual scene. One of the consequences of sole reliance on the criterion of original intent in rock art studies is that it is misapplied to the definition of a scene. I stress that my concern is with definition, not, for example, content. A perceptual analysis, as adopted here, has to be very wary of comment which might rest on culture-specific knowledge. So perceptual universals at best supply speculative or indirect information regarding the content of a depicted scene. But what about the definition of that same scene *as* a scene? Do we require an appeal to original intention in order to judge *that* a picture, rock art or any other, depicts an event, that is, constitutes a scene? Faced with the massive database left behind by the extraordinary recorder of southern African rock art, Harald Pager, and wondering how to differentiate scenes, presumably for the purpose of statistics, Lenssen-Erz [101] comes up with a list of criteria for scene identification, some relating to perceptual reception, some to original intention. But we cannot afford to run the two together in this way: *that* X is a scene and that its makers *intended* a scene are separate questions. To judge that it was so intended we need to ask the makers. If they are available, in the case of a contemporary picture, they may be surprised and reply “can't you *see*?”. This is because judging that it *is* a scene is not a matter of

historical reconstruction, as would be the case with determining the, inevitably culture-specific, *content* of the picture. It is simply a matter of *looking*. The giveaway is that rock art scholars who ask “did the makers intend this as a scene?” have already registered it as such—give or take borderline cases—and merely wonder if their seeing corresponds to that of the makers. Of course if they are analytically minded they may ask whether what they do *not* see as a scene *was* seen as such by its makers. But this question is as odd as the other, since it confuses seeing a scene with *knowing* the story behind the picture when it is not actually depicted as a story *in* the picture!

The argument that a scene is a scene if it was so intended has it the wrong way around. Certainly if a composition is perceived as a scene the chances must be that it was intended, otherwise it would be too much of a formal coincidence. In fact the present author has gone out of his way to concoct the scenario of a demonstrably unintended scene (Dobrez [1,97]). A non-concocted example comes to mind from Potash Road, Moab, Utah. It appears the original scene is that of hunters after sheep, over which is superimposed a large bear such as to give the impression that the *bear* is the object of the hunters’ attention. So the “bear hunt” could be read as an unintended scene—unless the second artist intended to change a sheep hunt into a bear hunt, in which case we have two intentions. Of course this is unprovable. All we can say is that we see a prominent bear hunt which, on closer inspection, may be a superimposition over a sheep hunt! So unintended scenes require unusual or at any rate particular circumstances. Normally a scene is not seen as such *because it was intended*. Rather it is likely to have been intended *because it is seen as a scene*. But why is it odd to wonder if our seeing corresponds to past seeing? Partly because the idea has no application beyond questioning all modern readings. After which nothing can be said about the picture under discussion that is necessarily about the picture under discussion. But the main reason for suggesting that it is odd to query perceptual correspondence between past and present is the absence of evidence for significant change in our visual system over a substantial period of time. The brain changed size some half a million years ago in the transition from *Homo erectus* to *neanderthalensis*. But experiments with monkeys, for which ample evidence has just been given, shows that we have very ancient and relatively stable visual structures.

I would surmise that, if challenged, not many rock art researchers will seriously propose that makers of rock art, in various places over varied time, actually saw in a way different from ours. More likely they will run together seeing (biological universals) and knowing (cultural specifics), that is, the phenomenon of visually registering “something happening” and the precise interpretation of that “something happening”, including its possible symbolism. Lewis-Williams (1986 [104]), in an article quoted by Lenssen-Erz, seems to do both. In fairness it should be said that the argument rightly castigates literalist Eurocentric readings of rock art. Lewis-Williams gives the example of a Drakensberg composition which features figures crossing what looks like, but is unlikely to be, a rope bridge. Given San-related ethnography, it may more plausibly depict figures “crossing” in a symbolic sense over a symbolic bridge. But at one point we have the following statement: “It is possible that no fixed relationships, as understood by Westerners, are intended in San ‘scenes’: the ‘scenes’ may portray figures individually and even in different temporal dimensions and therefore not capture an instant, in the manner of a photograph, a peculiarly Western concept” (p. 176). It may be best to leave the reference to post-Renaissance Western perspectival pictorial conventions, which eventually generated a mechanical equivalent in the camera, out of this discussion, since there is no necessary connection between event depiction (which indeed captures an instant) and perspectival depiction

(which, either via the post-Renaissance brush or the camera, captures it in a peculiarly Western way). Lewis-Williams' reference to "temporal dimension" I take to be saying that parts or all of a picture may be read symbolically, which is beyond doubt. Where a difficulty introduces itself is in the comment suggesting that San compositions that look like scenes might simply be a collection of separate figures not necessarily interacting or even acting. But this would make nonsense of the author's own readings of possible shamanic scenes, not least in this article. A scene is registered via a percept, not a concept; it is not inferred. We *see* a crossing and *infer*, via ethnographic information, its symbolic intention. Lewis-Williams' mention of Hedges' work on phosphenes and his own gloss that the "central nervous system is common to all men" (p. 174) might have prompted him to take as universalist a line in connection with scenes as he does in connection with phenomena of trance.

Still, Lewis-Williams' comment is made as a tentative afterthought. There are rock art commentators who (depending on how they are understood) appear to put forward a hard line on the issue of present reception of past images. Though it involves a digression from the consideration of depicted scenes, this sort of argument is well worth mentioning, since it not only bears on the intention/reception matters raised above, but also helps to clarify the implications of my own approach. Bednarik has pointed up some of the absurdities of rock art scholarship, concluding, in his all-or-nothing way, that readings of rock art images tell us nothing about the images and everything about the readers (2007 [105], p. 158). At the same time he has maintained that the sole criterion for a valid reading of rock art must be access to the intention of the makers ([105], p. 153). Since he believes such access barely exists, his position is logically consistent. However, it overstates the (as I have stressed, real) difficulties of the hermeneutic situation, that is, of recuperating past intentions, and this while equally rejecting reception as a valid avenue for understanding. To what extent is this last to be taken as scepticism about the constancy of perception itself? Certainly what is under attack in chapter 8 of Bednarik's *Rock Art Science* is not exactly our perception of rock art, but our attribution of content, that is, iconographic meanings to it—though there would seem to be some slippage between the two positions.

Elsewhere Bednarik specifically argues that moderns perceive differently from their forbears. This, as I understand, on the basis of comments by Helvenston (2013 [106,107]), who gives an account of well-known and compelling work on differences between the mindset of oral and literate cultures by scholars like Ong (1982 [108]). Ong's concern is with "consciousness" or "mind" rather than the brain, but Helvenston, while following his terminology, is clearly thinking in neural terms. It is equally clear from the definition she gives of her terms that her notion of a "modern mind" relates to the neuroplasticity thesis and not to evolutionary natural selection, that is, genetic transmission. In other words she is talking about wholly similar modern and ancient brains and the specific historical situation in which literacy might (modestly) reshape the brain through cultural transmission. In this scenario it is easy to understand how it is that a non-literate individual or group can in a short time pick up literacy. Though its effects could be significant, even (in theory) unlimited, in practice neuroplasticity does not affect the neural basics. For a start, operating as it does within the boundaries of an individual lifetime, it does not have time to do so. And in any case it would make no sense for a change in neural basics as long as the laws of physics governing our visual world remain unchanged. So I do not see Helvenston's comments as bearing on perceptual universals and therefore on the way we moderns actually *see* an ancient rock art image. Nor do I see how a possible appeal to *epigenetic*

transmission might alter the picture. Helvenston's point appears less radical than Donald's apparently modest proposal (1991 [109]) when he outlines his concept of modern "Theoretic" culture as involving reconfiguration of cognitive structures "without major genetic change" (p. 382). And in this context I am reminded of a comment by Ramachandran [110] who cautions that a Martian anthropologist observing earthlings might conclude that only genetic change could possibly explain the scientific/industrial revolution from the nineteenth-century on. But we know better, because we know the cultural serendipitous factors that have come into play.

To what extent Bednarik's position is more radical than Helvenston's depends on whether we take him as referring to the neuroplasticity thesis and how exactly he understands it. Certainly when reviewing the "earliest evidence of palaeoart" (2003 [111]) he assumes, tentatively yet inescapably, that anthropomorphic-looking stones (arguably artificially modified) like the Berekhat Ram (Golan Heights) or Tan-Tan (Morocco) ones would have been recognized as such in the exceedingly remote past. If that is not the case, then there is no basis for their candidacy as palaeoart. This even more so for the Makapansgat cobble (South Africa) in which we moderns discern the likeness of a face and which rates as *objet trouvé* palaeoart solely on the assumption that australopithecines two to three million years ago read it as a face. Like all of us, Bednarik is of course utterly seduced by the implications of such finds. So when he says that ancient artists and modern viewers do not "share the same brain, perception, graphic convention or belief system" (2013 [112]), p. 211, I agree entirely as regards graphic conventions and beliefs but favour a soft rather than hard interpretation of his comment as regards perception. Not least in the light of the vast amount of experimental work alluded to above relating, homologously but very closely, monkey to human visual systems, there would seem to be no evidence to support a view that even the most ancient rock art, let alone art made at the relatively recent time of transition between orality and literacy, is the product of humans with perception unlike ours. Recent culturally-driven change may have altered, and may still be altering, some of the visual system's circuitry, but not its hardwiring.

To return to the issue of intention and reception. We have easy, if partial and uncertain, access to the recent past (we know *something* of what *may* plausibly have been on Michelangelo's mind at the time of painting). Intention in the remote past presents another order of difficulty. But this does not mean its mindset is purely and simply closed to us. At the very least, the evidence, in the form of depictions, is there to *see*. That evidence, however complexly and deviously related to an original mindset, is available to (ideally, reflexive and critical) reception at a level common to all humans.

20. Fundamental Elements of Narrative Depiction

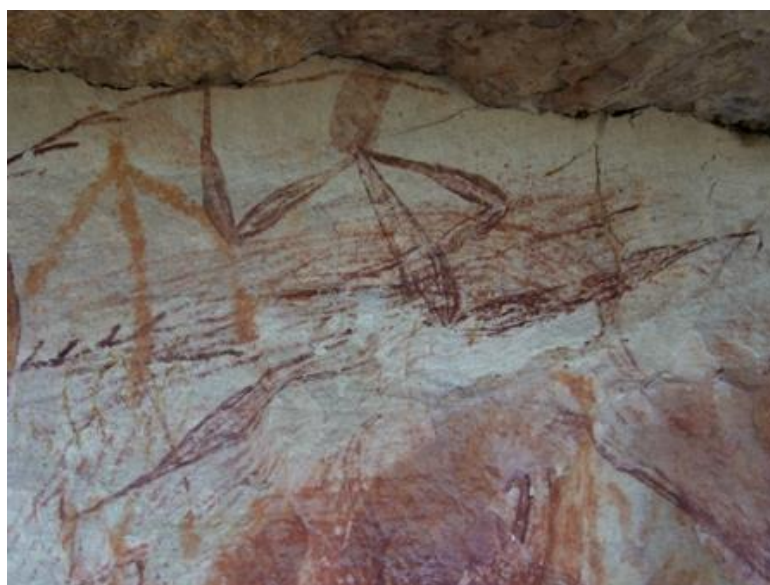
Focussing entirely on perception understood, in the light of the foregoing discussion, as a neural constant, disposes of fallacies and misapprehensions frequently advocated in rock art studies and enables us to make a list of elements which might reasonably contribute to the judgement "that X is a depicted narrative" (or event or scene or, in my terminology, a Narrative Depiction). The first of these elements must be movement, since without it there can be no visual narrative. Motion in a figure is that *sine qua non* which prompts the perception of an event, a "happening". Something like that idea was presumably on Leroi-Gourhan's mind when he linked motion and time and called the result "animation" (1982 [113]). However, what generates the perception of motion in a pictured figure is

asymmetry, what I have previously termed “markers of movement”. These are not Friedman and Stevenson’s pictorial movement “indicators” mentioned above and which mostly obscure the issues under discussion here. They are primarily bent or angled limbs or in one way or another asymmetrical figural compositions. Johansson’s experiments in which lights work when placed at body joints, or even between joints (Dittrich [6]), provide scientific evidence complementing these comments. This point need not be laboured, since so much of the above discussion shows unambiguously that perception of motion is a universal phenomenon (with pictorial motion-from-form, that is, movement registered via formal markers, processed, as argued, in the dorsal visual pathway). In some cases, even what at first sight appears to be a fortuitously widespread pictorial convention may suggest a hardwired perceptual mechanism. An example being, as in the Spanish Levant, southern Africa and Australia, a runner with characteristically-spaced legs forming a straight line at a tangent to the torso, that is, an extreme emphasizing of the torso-legs joint indicative of maximal speed (Figures 18 and 19):

Figure 18. Game Pass Shelter, Kwazulu Natal, South Africa.



Figure 19. Paperbark Beds site, Mt Borradaile, Arnhem Land, Australia.



There is another aspect of depicted motion across a visual field, namely that asymmetry or imbalance is transferred to the representational space of the picture. Thus a runner, as in the examples

just given, affects the observer's perception of the space around it: we see not only the figure as dynamic, but equally the space in which it moves as dynamic. In particular such space is seen as expanding behind the figure and contracting in front of it. Indeed, without stretching representational space in this way we do not register the impression of movement. Gibson noted that even with motion in an empty sky (as when we watch a bird in flight), space is not inert: "the *interspaces* between the edges of an object, and the nearest [other] edges constitute patches that are decremented and incremented" ([4] p. 103). It seems to me that Freyd's "implicit motion" (and RM) experiments effectively make this point for depicted motion: in overestimating the position of the jumping man we alter the space in which he moves, diminishing space between him and the ground and expanding space above him—and all this in a context which is actually *measurable*.

Is it necessary to add that in saying this I make no appeal to post-Renaissance pictorial conventions of foreshortening or perspective—these being simply *one* way of representing space in a picture? (For an account which puts European perspectival projection in proper perspective see Hagen 1986 [114].) In general rock art does not depict space directly. It certainly does not need to do so to evoke the sense of movement or event—and this alone challenges any idea that post-Renaissance perspective has a necessary connection with the creation of a scene. At the same time, rock art is perfectly able to depict space *indirectly*, not as an empty vessel which might be filled with Italian Renaissance streets and buildings, as in a painting by Piero della Francesca, but as dynamic. The famous Valltorta "hunt" shows this as well as any other rock art scene (Figure 20).

Figure 20. Cova dels Cavalls (model), Museu de la Valltorta, Tirig, Spain.



It allows a deal of empty space behind and among the cervids rushing from the right of the panel, and diminishes the space between hunters (on the left) and prey (on the right) to the point where they almost collide.

This example allows us to proceed from the depiction of a single-figure event, that is, a single figure in motion, as in Figure 19, to the depiction of movement involving a number of figures, either

moving together or placed oppositionally. Interaction, as I have argued, is not a category separate from action. It is merely a variation, either with one figure (as in the example of the “wounded man”) or more than one figure (the Valltorta “hunt”). I intend to say more about single-figure interaction below, but at this stage want to focus on interaction between two or more figures, where, as in all motion depiction, it is signaled by asymmetry, but in this case at the level of a more complex composition. We see the Valltorta event as such precisely because the asymmetries of angled limbs, outstretched cervid legs, *etc.*, have their counterpart in the asymmetry of the entire composition, including the space implied in it—which results in represented disequilibrium, figures tumbling from right to left only to come up against an obstacle in the line of hunters (at which point the observer anticipates the obstacle, space is compromised, and deceleration takes over). The whole picture would seem ready-made for an RM experiment!

Naturally these effects depend to a great extent on the distances between different participants in the depicted story. These are measurable and so could be tested in experiment—though only in given cases and not, as it were, across the board, since each case is affected by a unique variety of formal/perceptual factors, all mutually interacting and so complicating the result. At any rate it is evident that at some point too great a gap between participants abolishes perception of an integrated event. This is probably so with some (not all) panels of the so-called “swimmers” in the Australian Kimberley Mitchell plateau. The “swimmers” are more plausibly groups exchanging spears, but in some instances their placing means that it is easier to see them as separate rather than related. In addition to single- and multi-figure asymmetries, including asymmetries of representational space, as well as spatial separation of figures, there are other fundamental elements of Narrative Depiction to which I have alluded in the past and which require only brief mention here so as to relate them to the experimental work which is the focus of the present argument.

One is the orientation of figures, which has to be in profile. We saw that Johansson point-light displays require something approximating profile for recognition of form-from-motion. Ditttrich [6] specifically states that he mainly used point-light figures in side view. In depiction, the reason for profile is that it allows for interaction between figures. But more profoundly, profile depiction has the effect of *excluding* the viewer. This makes possible the situation I have termed “arm’s length event” perception. Only in this perceptual situation, whether in life, as in Johansson’s displays, or in depiction, as in Freyd’s jumping-man experiment and, equally, a variety of pictures, including rock art ones, is it possible to view an event without personal risk, that is, without risk to oneself as viewer. All narratives, not merely pictorially but in literature too, depend on this principle: that we register without involvement. Curiously, the size of figures in a picture has a significant role in a scenario of this kind. I have always noted that Narrative Depiction, that is, event depictions or scenes in rock art, tend to feature small figures, sometimes very small figures indeed, as in the case of the Valltorta ones. In the present context we can say that miniature representation helps to distance the viewer from the scene and so underlines the phenomenon of viewer exclusion. It is something we encounter in miniature traditions from Europe to Persia to India and which perhaps accounts in part for the miniature’s charm and its capacity to evoke a remote fantasy world.

21. The Perception of Causality

To see “something happening” in what I term a Narrative Depiction we need to register a mix—in variable proportions—of formal/perceptual elements listed so far. In previous work I have assumed this to suffice. However, there is more to it than that, and to focus on this “more” we must ask, with still greater precision, what it is we see when we see an interaction. It is not just a matter of additional visual complexity. Returning to Lenssen-Erz, we recall his need to distinguish different scenes. To this end he postulates factors which might “bind” figures together and in so doing constitute a scene. Of course in the process he adopts what I have ruled out as the Interaction Fallacy, the idea of scene defined solely as interaction of two or more figures. Moreover I do not regard his notion of “coherence” as adequate to the task of defining a scene, though I am encouraged by it to look for another candidate. Now what constitutes a *single* figure in action as a scene is certainly the perception of *coherent motion*, the sort of motion recognized by Johansson’s observers as indicating form. This matter requires no further elaboration. But what about *interaction* in a depicted event (whether involving a single figure, like the “wounded man”, or more than one figure)? What is it that generates that specific form of coherence in such a picture—over and above its coherent motion?

For events involving a number of interacting figures it is not merely the Gestalt “law of common fate”, first identified by Wertheimer and appealed to by Johansson [7]. Wertheimer’s law (clearly outlined in Köhler 1947 [115]) simply states that objects are likely to be visually organized into groups. As ever, the brain is prone to cutting corners: it is easier to register a “forest” than to add up individual trees, because the law of common fate is good for survival. But to register real or depicted figures doing something, either in common or in opposition, is to register more than a formal “glue”—and in any case “binding” is not exactly the issue with that tantalizing, yet to be explained, single “wounded” figure. Rather: *to see figures doing something interactively is to see causes and effects*. But is it possible to *see* (real or depicted) causal relations? The fact is we appear to do it whenever we read an event as an event. When we see X throwing a spear at an animal we attribute agency and when we see the animal wounded we attribute causality. We see X as cause and the wounding as effect. This would explain why it is we perceive the single “wounded man” as an interaction: it is an apparent effect with its cause as it were out of sight. The “wounded man” is someone to whom something has been done. But I intend to concentrate on scenes in which various figures interact. As it happens the phenomenon of causality has been investigated by the psychologist Michotte, whose experiments convincingly point to its direct perception. So not only do we directly see movement; we directly see interactions, that is, causal relations. This is not the place to restate Hume’s celebrated case against proof of causality. Such proof is not required when, on an everyday basis, we read causes and effects so as to remain in the evolutionary race. When you tap a billiard ball with a cue and the ball moves, you do not simply register a cue-motion/ball-motion “association”, then *infer* causality. If you did, you might as easily ascribe the motion of the ball to some other associational factor, say the glass of whisky on the edge of the billiard table. But no one infers such a relation, and this because you do *not* see the glass of whisky causing motion of the ball, but do indeed *see* the cue doing so.

Michotte put the direct perception argument in his 1946 (1963 [116]) *The Perception of Causality* and in a number of later essays published in English under the title *Michotte’s Experimental Phenomenology of Perception* (1991 [117]). He used simple equipment (a rotated disc giving an

impression of movement) to evoke responses in subjects under variable spatiotemporal conditions. The experiments *were* that, in the sense of being repeatable and controlled. At the same time the method was “phenomenological” in a broadly Husserlian sense in that it took account of response without dismissing it as subjective, while setting up conditions to counteract the *merely* subjective (by “concomitant variation”). Michotte acknowledged the truth of Hume’s (in practice, untenable) proposition that causality is unprovable, countering with arguments from Aristotle to Malebranche to Koffka so as to lead to the case of his own “experimental phenomenology”. With reference to a hammer driving a nail or a knife cutting bread he asks: “when we observe these operations, is our perception limited to the impression of two movements spatially and temporally co-ordinated.... Or rather do we directly perceive action as such—do we see the knife actually cut the bread?” ([116], p. 15). Two type-experiments focussed on what was termed the Launching Effect (*l’effet lancement*) in which we observe A bump B so as to push it forward, and the Entraining Effect (*l’effet entraînement*) in which A carries B along. Michotte argues that in these instances it is not a question of observers “interpreting” what they see, *since a strictly specifiable time interval suffices to cancel out the impression of cause and effect*. Rather causality is *itself* perceived. Relevantly for the present distinction between action and interaction, causality is not synonymous with movement perception, though it requires it—because movement perception is possible without perception of causality. In short, it is the perception of interaction over and above action that entails perception of causality.

Michotte’s work was appreciated by Gibson as very much in line with his own [4]. It is acknowledged by Ditttrich in his discussion of high-level categorial perception: “it seems evident that the perception of the goals of actions and the structures of action is not separated and then recombined at later stages of information processing. Thus, intentional or causal interpretations of motion displays can be perceived immediately” ([6], p. 21). For present purposes, however, we need to link the Michotte argument to perception of still pictures. A way of doing this suggests itself by reference to Hubbard *et al.* (2001 [118]), who tested the Launching Effect for RM-type displacement and came up with a positive result: “launching”, that is, *causality* (and not simply the movement involved) could be measured along the lines proposed by Freyd. Now Michotte, as explained, worked with illusory movement generated by a disc, that is, with a simulation of real movement, and the Hubbard *et al.* experiment involved real movement—each providing experimental support (of a different kind) for the everyday observation that we *see* causal relations. I would conjecture, though with confidence, that a Hubbard “displacement” experiment using *still images* would result in similar positive findings, so indicating that Michotte’s thesis applies to *depicted* causality. That would provide experimental support for the everyday belief that we *see* interaction in a depicted scene. Such a line of argument would account for the perceptual “binding” of figures in a scene which Lenssen-Erz was seeking. What binds, that is, makes an interactive scene, is the direct perception of cause and effect—along with perception of associated phenomena, such as agency. This cannot apply to scenes of action alone (say, “man running”), there being no issue of causality involved. But it can certainly apply to implied interaction (the “wounded man”). In the case of action events, motion, provided it is not random but whole or “coherent”, suffices to prompt the impression of “something happening” (“man running”), but goes no further. At any rate, with respect to interactions, the brain privileges large perceptual wholes over atomistic units of the sort presumed in Hume’s thinking, such that interpretation of causality is not as it were “added on” to the percept but built into it from the start. Neural systems are

more likely to feed-forward, or fast-track, reading of an event by referring early visual processing to high-level areas.

Where might such areas for the processing of perceived cause and effect be located? On the basis of earlier sections of this paper, we might predict the upper reaches of the dorsal motion pathway—and this would in fact seem to be the case. Pelphrey *et al.* [73] carried out a point-light experiment involving human responses to biological motion and found response in the human STS which suggested sensitivity to the *social* dimension of movement, that is, readings of *intention* (even in an animated robot). Peeters *et al.* (2009 [119]), using monkeys and humans responding to a video featuring activities with tools, found activation in the occipitotemporal, intraparietal and premotor areas, in short, and expectedly, in the higher dorsal path. Most specifically, there was a *human-only* response in an area (labelled aSMG) of the left inferior parietal lobule (IPL), which led the researchers to postulate a coding for *causal relations*. If it is the case that the higher sections of the dorsal motion processing path actually analyze perception of causality in a situation of real motion (a video display) we may, going by the convergence of real and depicted motion perception discussed above, reasonably predict that perception of causality in a still picture would also be located in the same upper-dorsal pathway areas. In addition, it is probably the case that perception of cause and effect depends on the activity of mirror neurons. Thus it is in the mirror-neuron system that we might expect to uncover the precise neurophysiology of Michotte's results. I do not intend to go into the question of mirror neurons here, but for relevant perspectives see Patricia Dobrez in the present issue of *Arts*.

22. Causal Agency Relations in Narrative Depiction

It is not difficult to see how Michotte's experiments on causal interactions, extended to depicted motion, might have application to the perceptual analysis of pictures. I have pointed out that there is a critical *distance* between figures in a composition required for perception of interaction, in other words, cause and effect. This may be identified with that specifiable time interval required for an impression of causality in Michotte's experiments. Thus a *time* interval in a real situation translates into a *spatial* interval in a picture. (Of course this would not apply to implicit or single-figure interactions, namely the "wounded man".)

As regards figure orientation we may note that Michotte's artificial visuals for establishing, for example, the Launching Effect, correspond strictly to relations specified by *profile* depiction. The Madhya Pradesh boar (Figure 5) chasing a human is at once read as intending to be the agent of the man's death, that is, as an event involving clear cause and (predicted) effect. Likewise with this rhinoceros from Namibia, which is read as scattering a group of humans (Figure 21). Again the relation between figures is entirely causal and that is just what is meant when rock art researchers intuitively register it as a case of interaction. The same is true for the reading of the Valltorta composition (Figure 20). It is not even strictly necessary for us to see it as a *particular* kind of interaction, namely a "hunt", though that is of course likely. It suffices to see asymmetries at the level of figures and of the composition as a whole, coupled with the fact of profile depiction and judgement regarding spatial intervals, for us to register causal relations: cervids and humans are on a collision course. Other compositional combinations work to the same end. A line of tiny humans on the left of a Pressa Canyon, Texas, picture interact not with each other (their action being action-in-common) but,

as indicated by orientation, with the group on the right, themselves focussedly interacting with a single central figure (Figure 22). The result may be a reading of these combinations as a “birthing” event, though this is not necessary for the point I am making. Something similar may be said of a detail from a Serra da Capivara composition in Brazil, also referred to as a “birthing” scene (Figure 23).

Figure 21. Noukloof Mountains, Namibia.



Figure 22. Pressa Canyon, Texas, USA.

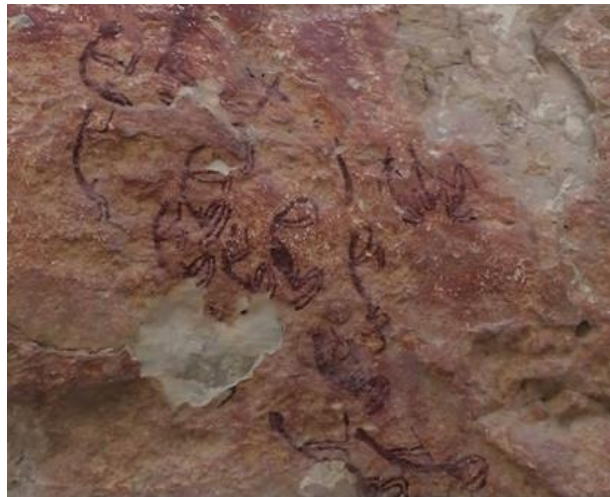


Figure 23. Desfiladeiro, Serra da Capivara, Piauí, Brazil.



The Writing-On-Stone scene illustrated earlier (Figure 6) involves a reading of multiple and highly complex causal relations. Following Keyser and Klassen's (2001 [120]) description and the plausible argument that what is depicted here is an actual event which occurred in 1866 on the Milk River, we have the following: on the right, figures advance, weapons pointing, towards a central episode involving casualties; on the left, figures look to be halted in front of a tipi encampment by a line of firing rifles, some held by shooters, others free-standing. There are tracks indicating movements, and dots indicating bullets passing between antagonists. Via ethnographic evidence, this may be presumed to be a battle in which a party of Gros Ventres, Plains Cree and Crow attack and kill two Piegans (in the centre of the composition), then advance (right to left in the composition) to the Piegan camp where, however, they are defeated (by that overwhelming lineup of rifles). But even without ethnographic help the observer registers the complex interplay of causality: threatening intentional movement (of course rendered through profile orientation) on the right, culminating in specific interaction, for example an axe-wielding figure shot by a figure placed in opposition to it, in the middle—at which point space between antagonists is minimized; and on the left minimal space interval between rifles and opposed group, all of it highlighting interaction. The interesting thing about these kinds of rock art compositions in the northwestern USA and southern Alberta is that they do not necessarily emphasize markers of movement. There are legs open as for “walking”, arms raised as for “striking”, *etc.*, in the Writing-On-Stone battle scene, but much of the action is represented in shorthand, with weapons indicating those using them, dot conventions indicating, via “tracks”, movement from A to B, a convention of depicting flying bullets, *etc.* Whatever requires inference or knowledge of pictorial conventions is of course matter for something other than direct perception. With a picture of this kind we are on the borderline of Narrative Depiction and symbolic depiction, the one a matter of *seeing*, the other of *knowing-about*. By this logic symbols cannot depict an event directly; instead, they tell us about it, provided we know the semiotic code. They cannot be universals. Nonetheless, the Writing-On-Stone panel functions visually as a scene, albeit modified by a culture-specific sign system. We see “something happening” and what structures its detailed local events as well as structuring the whole as an integrated complex event is the fact that the observer immediately registers interaction as causal.

So an interactive scene is integrated by perceived causality. But we cannot define scene as such in these terms. Scenes or Narrative Depictions are constituted by those fundamental representational elements listed above and characterized as event perceptions of an “arm's length”, non-involved kind—what I have set in the situation in which the observer of the narrative or scene is not at risk, but simply *looks on*. Of course Narrative Depictions are not restricted to rock art, though this paper has stressed it. Depicted events featuring single or many figures recur in Ukiyo-e (Japanese wood-block narrating sexual, military and supernatural episodes); throughout Asia they recur as frescoes on temple walls recounting the life and past lives of the Buddha; in India as Mughal miniatures or miniatures of Krishna, either amorous or mischief-making; in Persia as the Shah-nama or Book of Kings; in Christian Europe as repeated episodes in the life of Christ, his mother, her parents, *etc.*: Annunciation all the way to Crucifixion, Resurrection, *etc.*; in secular Europe as, to take one example, those extravagant canvases of Napoleonic battles by Gros, Lejeune and others. In all cases, not least depictions of horrendous slaughter and mayhem, the observer looks on in complete security, excluded by factors which precisely define the type of perception appropriate for a scene.

23. Fundamental Elements of Performative Depiction

So much for perception of narrative events in life and art. There is, of course, that other type of event perception, the one involving risk for the observer, namely that of visual “looming”. I have said that point-light experiments with biological motion prefer figures in profile, but this is the place to add that they actually exclude, and for good reason, movement *towards* the observer (Dittrich [6]). Such displays demand unambiguous views of joints in the human body for the purpose of recognition, namely that “X is walking, running”, *etc.* However, the looming effect precludes attention to body joints, that is, detached observation of what X is doing—the kind of observation that would be appropriate for a scene. Rather looming engages attention on what X is doing relative to the observer. Movement markers which operate in this situation are reduced to one: the progressive *expansion* of X as it approaches the observer. We know from the Kelly and Freyd [87] experiment earlier described that looming prompts RM (or at the very least, visual displacement), that is, that “growing” figures—expanding in size and so taken as “approaching”—are remembered as larger than they in fact are. Given that RM experiments (mostly) use static images, we are experimentally, as well as experientially, justified in concluding that the looming effect applies to pictures. But how *exactly* does it translate to the situation of depicted motion?

Any pop visual text such as a cartoon, comic or graphic novel, illustrates the point. Indiana Jones (Figure 24) breaks out of a magazine page frame to deliver his punch—right in the observer’s face!

Figure 24. Indiana Jones, archaeologist.



The visual effect depends on there being a frame and, equally, on post-Renaissance perspectival conventions, in this case extreme foreshortening: the fist projects outwards, in your face. At the same time eye-contact ensures a reading of agency or intent: he’s looking at *me*—and he’s aiming for *me*! All this realizes Michotte’s situation of the direct perception of causality. Indeed the look alone is sufficient to generate a looming effect (something I suggest would be worth an experiment). If a figure in profile looks at another figure, then we see a scene involving interaction and which does not involve

us; if the figure is facing us and looking at us, the impression of a scene cannot be sustained because what is established is a direct relation between the figure and the observer, that is, a looming effect. Actually the look on its own is enough to establish this coming-towards. Manet's "*Le Déjeuner sur l'herbe*" (1863) in the Musée d'Orsay depicts a picnic scene which includes three foregrounded figures: a naked female and two fully-clothed males, these last engaged in conversation. However, there is a complication: one of the figures—the female nude—looks straight out of the picture at the viewer. This at once establishes a connection which is equivalent to Indiana Jones coming out of the page-frame. The same effect obtains with the "*Olympia*" (1863, Musée d'Orsay), Manet's picture of a naked prostitute with her servant about to hand her a bunch of flowers. It would constitute a scene—except that the prostitute is looking *out* of the picture, that is, making eye-contact with the viewer, now put in the position of the customer who has made the present of flowers.

A sixteenth-century example analogous to the above is Parmigianino's self-portrait in a convex mirror in the Vienna Kunsthistorisches museum. The artist looms in the mirror, that is, looms at the viewer, simultaneously making eye-contact with himself in the mirror and you the viewer outside the picture—and extending his hand towards you (which is disproportionately large and therefore appears very close). Perhaps the best-known example of depicted looming is the World War I recruiting poster featuring Kitchener, which works so well it has been much copied and adapted (Figure 25).

Figure 25. "Your Country Needs You" by Alfred Leete, 1914, UK.



Kitchener looks out at you, his arm extends towards you, his finger points at you (and at no one else), and the message is: "Your country needs you!". In all such cases, what we perceive is a direct contact, the sort that, in real life, makes you take evasive action. Direct engagement in real life is risky and, even in pictures, it is disturbing or at the very least disconcerting. But does it really need to depend on perspectival conventions?

The answer is: clearly, *not*. Rock art rarely makes use of perspective of the sort familiar to post-Renaissance Europeans. European rock art scholars desperate for signs of what they (wrongly) take as pictorial sophistication have had only minor success in finding cases of foreshortening (as, for example, with the Lascaux leaping cow). But these are exceptions. What about the convention of picture-framing? Can you depict the phenomenon of "breaking out of the frame" if there is no frame? On the face of it rock art has no frame around it, but this assumption is actually mistaken. In fact all

depictions are framed in one way or another, either materially by something around or behind them (behind them in the form of a support, say a wall or rock face) or, more strictly and non-materially, by a perceptual *act* by which we focus attention on an object—the depiction—to the exclusion of what is *not* the object. The frame is really a matter of attention (and related to the neural substrates for attention), not a matter of material enclosure or limiting support. (For discussion of this see Dobrez 2010 [121].) Nonetheless examples given above do entail the presence of a material frame: a picture frame or the edges of a page or poster. Rock art has no overt framing of this sort and therefore cannot duplicate the perspectival tricks found in pop images, Parmigianino, the Kitchener poster, *etc.*

But all the same, it has perfectly identifiable ways of evoking the sense of looming, ways which long predate perspectival projection. We recall the fundamental elements listed above as constituting depicted events of the sort observed as it were “from a safe distance”: movement markers in the form of asymmetries, profile orientation, the generation of representational space, *etc.* To obtain a looming effect *without* perspectival projection, as in rock art, the movement marker required is precisely the one which mimics looming, that is, in real life, progressive increase in size. But rock art depictions are stills. So rock art frequently has primary recourse to size itself. It is an observable fact that rock art scenes worldwide tend to feature relatively small, sometimes tiny figures, and looming compositions tend to feature relatively large, sometimes very large figures. This may be relative to the observer’s size, or to that of other figures in the composition, or relative to the size of the support (say, covering an entire rock surface, provided the rock is not too big), or it may be a function of relations *within* the figure which give a sense of “monumentality”, as we find, for example, in late Raphael or Michelangelo—or images at Barrier Canyon, Utah (Figure 26).

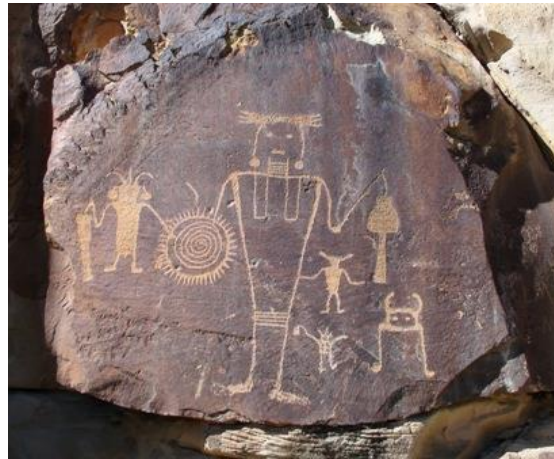
Figure 26. Barrier Canyon, Utah, USA.



Barrier figures, while small in relation to the rock cliff on which they appear, are large relative to the viewer and monumental in their internal proportions. They generally have no distracting limbs which

might generate asymmetry, that is, motion more appropriate to a scene. Above all, they are full-frontal. Symmetrical frontality plus size generates a strong sense of looming—and sometimes symmetrical frontality alone will do the job, as with sizeable but not overly large Vernal figures, also from Utah (Figure 27).

Figure 27. McKee Spring, Vernal, Utah, USA.



In formal terms we might think of such images as static, but this is far from being the case perceptually. It is just that movement is not tied to asymmetries such as bent limbs, and that frontality ensures that images, rather than interacting among themselves (as would be the case in a profile-depiction scene), interact with the viewer, that is, loom in the viewer's direction. This must be so because movement which is not lateral has nowhere to go except depth, which in this case means *out*. Looming figures are therefore read as intensely active: though they do nothing among themselves, their doing is directed outwards and this is frequently underlined by eye-contact: they often feature prominent eyes. In the case of Australian Wandjinas, whether painted on a rock or on bark, even the potential distraction of the mouth is removed, while monumental frontality is heightened by a halo or aura around the head (Figures 28 and 29):

Figure 28. Chamberlain Gorge, Kimberley, Australia.



Figure 29. Bark painting. Artist unknown, Kimberley, Australia.



Looming necessarily precludes observer neutrality. If a depiction looms at you it must generate a completely different representational space from that of a scene, indeed, in strict terms, it must generate something other than representational space. In a scene, as I have argued, there is an “inside” (the scene) and an “outside” (the observer), and the two do not touch. You cannot enter the representational space of a scene, since by definition a scene excludes you. But the opposite obtains with a looming image: it is perceived as *intrusive*. To say it looms at you is to say it generates the impression of entering *your* space, that is, of emerging from representational space so as, at least after a fashion, to advance into *real* space. Gibson understood this with respect to his looming experiment described above: the dark patch on the screen which expands in size is actually perceived as coming “*out of the screen*” ([4], my italics). Naturally a depicted image does not move, but the impression is analogous. It is yet another case of the visual system’s near-enough principle. We do not confuse real looming with depicted looming any more than real with depicted scene. Nonetheless just as we read a depicted scene as “something going on” (in that “inside” space), so we read, indeed *see*, that other depicted event of something coming out of its representational frame to as it were directly engage us. With or without material frame, the “breaking of the frame” mechanism *works*. In the light of the earlier discussion of neural substrates for the looming effect we may suppose with probable accuracy that the *depicted* looming effect is generated in those same areas of the visual system that process looming in real life. (See Krekelberg *et al.* [85] on radial “implied” motion registered in MT.) At the same time—and recalling an above remark in connection with the possible neurophysiology of causality—it is probably the case that mirror neurons play a key role in allowing us to read movement *towards* the observer in much the way they might play a role in allowing readings of cause and effect, that is, intention (or motivation or agency) in visual *narratives*. For an account which links the mirror-neuron thesis to looming performative images, with specific reference to hand marks, see Patricia Dobrez in the present issue of *Arts*. It should be added that, just as real looming activates strong limbic response (for example, fear), so, in depiction, looming may be expected to evoke a sense of dominance or authority, or even to overawe or terrify, not least because in rock art looming figures

may well be set higher on a rock face than the viewer's position. This is often true with American examples: the Vernal figures, the imposing panels at Barrier Canyon and Dinwoody Lake—this last illustrated at the beginning of this article (Figure 7). It is also frequently true of Australian Wandjinas, say those at Chamberlain Gorge (Figure 28), and with the so-called Milbrodale Baiame or spirit-being which makes eye contact with the observer from above while enfolding the entire landscape below with its giant outstretched arms (Figure 30).

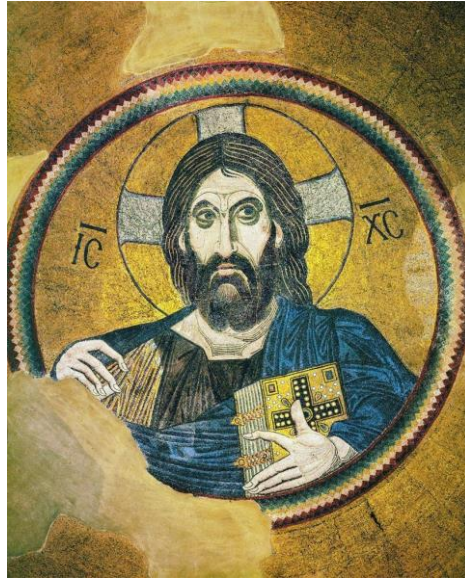
Figure 30. Milbrodale, Hunter region, New South Wales, Australia.



It is true that in the same area and further south, around Sydney, many comparable Baiames are pecked horizontally on the sandstone ground. In that case looming is, as so often, primarily a function of frontality and size.

Of course with these and other examples specific viewer-response is inevitably modulated by cultural factors. Response to a picture *as looming* is, I suggest, as universal as the urge to duck when the object in the 3D screen jumps at you. The precise feelings which colour such perception must be culture-specific. Still, it is notable that religious ikons often take the form of looming images, even if, as already made clear, looming images cannot be restricted to a particular pictorial category. A typically Byzantine Christ looks down (not quite at the observer) from the dome of the church at Dafni, Greece (Figure 31).

Figure 31. Pantocrator, Dafni, Greece. Courtesy Sasha Grishin.



The effect is as relevant to 3D figures, notably Buddhas, which may be overwhelmingly large (Figure 32).

Figure 32. Wat Suthat, Bangkok. Thailand.



Frontality in this example from Wat Suthat, Bangkok, is such that, though it is possible to walk around the figure, there is little point in doing so.

The Christian crucifix is more adaptable: it may be situated in a scene, so as to constitute a picture of a man being crucified, but often its frontality and position in architectural space ensures that, whether 2D or 3D, it will be seen as looming, less as an event to be observed from a distance and more

as a hieratic presence making direct contact. This is the case with the modest-sized Christ of S. Damiano (now in a side chapel of the basilica of St. Clare), Assisi, which celebratedly asked Francis to “repair his church”. This Christ, unlike many in the Italian tradition, does not look as if dying; rather his open eyes address the viewer (believer or not)—as does this crucifix from New Guinea (Figure 33).

Figure 33. Sepik region, New Guinea.



Unsurprisingly, even without ethnographic help, the rock art researcher will be liable to take any looming image as at least potentially relating to religious experience. It is difficult not to do so with Dinwoody-type images, like this one above Torrey Lake, Wyoming (Figure 34).

Figure 34. Torrey Lake, Wyoming, USA.



It should be evident now how I would like to distinguish “Performative” depictions like the above from Narratives or scenes. Narrative images tell a story either by way of depicted motion alone or depicted interactions involving causality. Narratives and Performatives are both registered as events (“something happening”), but they are signaled by a different, and in each case cohesive, set of features. Why “Performative”? I use the term as having associations with the Speech Act philosophy developed by Austin and Searle, where a statement is said to be “performative” if it effects what it says, like the “I do” of the marriage ceremony or the “I declare X open” of public occasions, both of which *do* what they *say*. In short, by the term Performative I mean an image which is active, but not after the manner of Narratives. A Performative represents motion, and so is as central to the present argument as a Narrative, but it represents motion in an unsettling direction—towards me, the observer. Such motion has never been properly noted, to my knowledge, let alone studied, yet it features in an entire range of depictions. In evolutionary terms it may plausibly be understood as perceptually more critical than the scene, since it engages directly in a way the other does not, that is, entails *risk* for the perceiver. Of course it is easier to accept the case for perception of *frontoparallel* movement in a picture, with its key markers of asymmetry/instability, than for perception of looming, that is, movement *in depth* in a picture. With real looming, objects increase in size as they approach. In a stationary Performative image of the sort illustrated above, frontal depiction and (relative) size have to do all the work, the only instability being *forward*, towards the observer—something less evident than lateral asymmetry as prompting the sense of an unstable and therefore *moving* object. And yet still pictures *do* indeed loom, observably so, and not merely in the more obvious cases which exploit perspectival depiction. Naturally, it is always possible to *resist* movement markers. Just as we can freeze a depiction of lateral movement (we recall Hockney’s leopard), so also we can refuse to “go along with” markers of looming in a picture. Unlike real motion, the static depiction cannot *force* perceptual response. What markers of movement, lateral or in depth, *can* do is to *encourage* a perceptual response. I want to argue that all this is phenomenologically observable. However, it is worth adding that it is also testable. Thus what I have said about two (primordial and continuing) types of depiction could be made the subject of experiments along Freyd lines. For an elaboration of this point I refer the reader to P. Dobrez on the subject of looming hand images in this issue of *Arts*.

24. A Universalist Approach to Depiction

This study of motion along lines suggested by evolution has covered not inconsiderable ground. Much of what I say in the way of analysis of basic pictorial types and the elements that go to making them has not been said before, nor has anyone as far as I know come up with an approach to the subject which is comparably inclusive, firmly tying depicted situations to real ones; taking in its scope *all* images, from rock art to forms more usually considered in Art History; and, finally, addressing a number of academic disciplines: the Humanities ones of philosophical hermeneutics and Art History, the Social Sciences ones of Archaeology and Anthropology currently responsible for most of the work in rock art studies, and the experimental sciences ones of cognitive psychology and neurophysiology.

In connection with Art History we may certainly say there have been examples of a biologically-oriented approach to representation, both from scientists and art historians (Zeki [81], Ramachandran 1999, 2003 [122,123], Livingstone [33], Onians 2007 [124]). But in general the

discipline is slow to move in this direction. At the same time only a minority of art historians take rock art seriously into account. Rock art researchers and commentators (to date mostly, though not solely, archaeologists and anthropologists) are rarely familiar with Art History, and are largely unfamiliar with biological approaches—this last with notable exceptions not inevitably comfortable in each other's company (Blackburn 1977, Hedges 1982, Bednarik 1984, Lewis-Williams 1988, Hodgson 2000, Watson 2009 [125–130]). On the other hand experimental psychologists and scientists have generally not had the confidence to tackle art, rock art included (with rare exceptions including those noted immediately above). Very few, if any, in the relevant disciplines are likely to take a philosophical line, and if they have done so (say, in Art History) it is probably without much, if any, reference to the immense field of rock art, whose time depth is greater than that of any other pictorial category.

In general what is lacking is something that will bring historically separate disciplines, with their more or less distinct methodologies, together, and it seems to me that focus on the phenomenon and mechanisms of perception is one way of fulfilling that need. Summers 2003 [131], who made a supreme effort to write a theoretically-savvy global history of art in the shadow of earlier histories and of post-structuralism, still faced the difficulty of finding a principle that might enable the enterprise. What to include in a global history, but, more to the point, on what *basis*? I suggest putting forward a plausibly available universal principle, based on the facts of *perception*, themselves based on evolutionary imperatives and the visual system these have brought into being. The combination of focus on *all* representation, especially including rock art, with methodological stress on perceptual universals, offers a way out of the Art Historical dilemma. Of course in saying this I do not imply that other approaches to the pictorial, including formal and/or historical ones, are somehow invalid. This is not the case, and analogous comments apply to rock art research as practised via culturally-oriented disciplines like Archaeology and Anthropology.

Moreover the kind of phenomenologically-guided analysis, with scientific backing, illustrated in this paper may well assist the scientist, in some instances suggesting to, for example, perceptual psychologists, specific experiments worth making. Read in the context of the preceding discussion, the following come to mind: does the subject perceive motion, or not, in a given picture? Does she perceive with or without post-Renaissance perspectival cues? Does her perception differ in the case of a scene, or a looming image? Does she register the effects investigated by Manet in some of his paintings? Does RM obtain when she views the Valltorta composition? How does she read causality in pictures? How does she register those fundamental elements of scene depiction (spatial separation, profile, *etc.*)? And those equally fundamental but different elements of looming depiction (not least the effect of the “look”)? It could also be that aspects of the above analysis might alert the neuroscientist to possible neural mechanisms worth investigation.

All these questions and suggestions are predicated on the view that understanding perceptual processes provides a key unlocking many answers. It seems we are made to see motion directly and to see a picture in much the same way as the real thing. The other point to add is a visual-typological one: that, through evolution, we actually see different kinds of pictures, different and hitherto unrecognized *categories* of pictures, very differently—and this in ways amenable to fundamental analysis.

Note

Copyright prevents reproduction in the present context of artworks by Hockney, Manet and Parmigianino mentioned above. Interested readers may access these online.

Acknowledgments

I am grateful to Jennifer Freyd for permission to use two of her illustrations, to James Keyser for permission to reproduce his tracing of the Writing-On-Stone “battle” scene, and to Sasha Grishin for the use of his photograph of the Dafni Pantocrator.

Conflicts of Interest

The author declares no conflict of interest.

References

1. Dobrez, L. Rock art, perception and the subject/object binary. *Rock Art Res.* **2011**, *28*, 71–83.
2. Gibson, J.J. *The Perception of the Visual World*; Houghton Mifflin: Boston, MA, USA, 1950.
3. Gibson, J.J. *The Senses Considered as Perceptual Systems*; Houghton Mifflin: Boston, MA, USA, 1966.
4. Gibson, J.J. *The Ecological Approach to Visual Perception*; Houghton Mifflin: Boston, MA, USA, 1979.
5. Mather, G.; West, S. Recognition of animal locomotion from dynamic point-light displays. *Perception* **1993**, *22*, 759–766.
6. Dittrich, W.H. Action categories and the perception of biological motion. *Perception* **1993**, *22*, 15–22.
7. Johansson, G. Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* **1973**, *14*, 201–211.
8. Johansson, G. Spatio-temporal differentiation and integration in visual motion perception. *Psychol. Res.* **1976**, *38*, 377–393.
9. Rosch, E.; Mervis, C.B.; Gray, W.D.; Johnson, D.M.; Boyes-Braem, P. Basic objects in natural categories. *Cogn. Psychol.* **1976**, *8*, 382–439.
10. Dobrez, L.; Dobrez, P. Rock art animals in profile: Visual recognition and the principles of canonical form. *Rock Art Res.* **2013**, *30*, 75–90.
11. Walk, R.D. Event Perception, Perceptual Organization and Emotion. In *Cognitive Processes in Perception and Art*; Crozier, W.R., Chapman, A.J. Eds.; Elsevier Science Publishers B.V.: Amsterdam, The Netherlands, 1984; pp. 211–221.
12. Runeson, S.; Frykholm, G. Visual perception of lifted weight. *J. Exp. Psychol.* **1981**, *7*, 733–740.
13. Bertenthal, B.I.; Proffitt, D.R.; Cutting, J.E. Infant sensitivity to figural coherence in biomechanical motions. *J. Exp. Child Psychol.* **1984**, *37*, 313–230.
14. Shepard, R.N. Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking and dreaming. *Psychol. Rev.* **1984**, *91*, 417–447.

15. Freyd, J.J. The mental representation of movement when static stimuli are viewed. *Percept. Psychophys.* **1983**, *33*, 575–581.
16. Pylyshyn, Z.W. The imagery debate: Analogue media *versus* tacit knowledge. *Psychol. Rev.* **1981**, *88*, 16–45.
17. Neisser, V. The Self Perceived. In *The Perceived Self: Ecological and Interpersonal Sources of Self-Knowledge*; Neisser, V., Ed.; Cambridge University Press: Cambridge, UK, 1993; pp. 3–24.
18. Freyd, J.J.; Finke, R.A. Representational momentum. *J. Exp. Psychol.* **1984**, *10*, 126–132.
19. Freyd, J.J. Dynamic mental representations. *Psychol. Rev.* **1987**, *94*, 427–438.
20. Schütz-Bosbach, S.; Prinz, W. Prospective coding in event representation. *Cogn. Process.* **2007**, *8*, 93–102.
21. Thornton, I.M.; Hayes, A.E. Anticipating action in complex scenes. *Vis. Cogn.* **2004**, *11*, 341–370.
22. Blätter, C.; Ferrari, V.; Didierjean, A.; Marmèche, E. Role of expertise and action in motion extrapolation from real road scenes. *Vis. Cogn.* **2012**, *20*, 988–1001.
23. Friedman, S.L.; Stevenson, M.B. Perception of Movement in Pictures. In *The Perception of Pictures, Volume 1: Alberti's Window: The Projective Model of Pictorial Information*; Hagen, M.A., Ed.; Academic Press: New York, NY, USA, 1980; pp. 225–255.
24. Hubbard, T.L. Approaches to Representational Momentum: Theories and Models. In *Space and Time in Perception and Action*; Nijhawan, R., Khurana, B., Eds.; Cambridge University Press: Cambridge, UK, 2010; pp. 338–365.
25. Hubbard, T.L.; Bharucha, J.J. Judged displacement in apparent vertical and horizontal motion. *Percept. Psychophys.* **1988**, *44*, 211–221.
26. Reed, C.L.; Vinson, N.G. Conceptual effects on representational momentum. *J. Exp. Psychol.* **1996**, *22*, 839–850.
27. Senior, C.; Foley, E. Cognitive neuroscience and (very) high level vision: The state of this union. *Optician* **2006**, *231*, 16–18.
28. Gibson, J.J. The information available in pictures. *Leonardo* **1971**, *4*, 27–35.
29. Kerzel, D. A matter of design: No representational momentum without predictability. *Vis. Cogn.* **2002**, *9*, 66–80.
30. Hubel, D.H.; Wiesel, T.N. Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **1959**, *148*, 574–591.
31. Hubel, D.H.; Wiesel, T.N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **1962**, *160*, 106–154.
32. Hubel, D.H.; Wiesel, T.N. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* **1968**, *195*, 215–243.
33. Livingstone, M. *Vision and Art: The Biology of Seeing*; Abrams: New York, NY, USA, 2002.
34. Ungerleider, L.G.; Mishkin, M. Two Cortical Visual Systems. In *Analysis of Visual Behavior*; Ingle, D.J., Goodale, M.A., Mansfield, R.J.W., Eds.; The Massachusetts Institute of Technology Press: Cambridge, MA, USA, 1982; pp. 549–586.
35. Dobrez, L. What representations tell us about the way we see (with Commentary by R. Bednarik and Reply). *Purakala: The Journal of Rock Art Society of India* **2010–2011**, *20–21*, 23–41.

36. Dobrez, L. American ikon: How to choose an ARARA logo. *Am. Indian Rock Art* **2012**, *38*, 145–164.
37. Livingstone, M.; Hubel, D. Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science* **1988**, *240*, 740–749.
38. Maunsell, J.H.R.; van Essen, D.C. Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *J. Neurophysiol.* **1983**, *49*, 1127–1147.
39. Shipp, S.; Zeki, S. The organization of connections between areas V5 and V1 in macaque monkey visual cortex. *Eur. J. Neurosci.* **1989**, *1*, 309–332.
40. Boussaoud, D.; Ungerleider, L.G.; Desimone, R. Pathways for motion analysis: Cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *J. Comp. Neurol.* **1990**, *296*, 462–495.
41. Goodale, M.A. Visual pathways supporting perception and action in the primate cerebral cortex. *Curr. Opin. Neurobiol.* **1993**, *3*, 578–585.
42. Allman, J.M.; Kaas, J.H. A representation of the visual field in the caudal third of the middle temporal gyrus of the owl monkey (*Aotus trivirgatus*). *Brain Res.* **1971**, *31*, 85–105.
43. Gattass, R.; Gross, C.G. Visual topography of striate projection zone (MT) in posterior superior temporal sulcus of the macaque. *J. Neurophysiol.* **1981**, *46*, 621–638.
44. Zeki, S.M. Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J. Physiol.* **1974**, *236*, 549–573.
45. Weller, R.E.; Kaas, J.H. Retinotopic patterns of connections of area 17 with visual areas V-II and MT in macaque monkeys. *J. Comp. Neurol.* **1983**, *220*, 253–279.
46. Ungerleider, L.G.; Desimone, R. Cortical connections of visual area MT in the macaque. *J. Comp. Neurol.* **1986**, *248*, 190–222.
47. Maunsell, J.H.R.; van Essen, D.C. Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *J. Neurophysiol.* **1983**, *49*, 1148–1167.
48. Maunsell, J.H.R.; van Essen, D.C. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J. Neurosci.* **1983**, *3*, 2563–2586.
49. Albright, T.D. Direction and orientation selectivity of neurons in visual area MT of the macaque. *J. Neurophysiol.* **1984**, *52*, 1106–1130.
50. Newsome, W.T.; Paré, E.B. A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J. Neurosci.* **1988**, *8*, 2201–2211.
51. Colby, C.L.; Gattass, R.; Olson, C.R.; Gross, C.G. Topographical organization of cortical afferents to extrastriate area PO in the macaque: A dual tracer study. *J. Comp. Neurol.* **1988**, *269*, 392–413.
52. Saito, H.; Yuki, M.; Tanaka, K.; Hikosaka, K.; Fukuda, Y.; Iwai, E. Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *J. Neurosci.* **1986**, *6*, 145–157.
53. Leinonen, L.; Hyvärinen, J.; Nyman, G.; Linnankoski, I. I. Functional properties of neurons in lateral part of associative area 7 in awake monkeys. *Exp. Brain Res.* **1979**, *34*, 299–320.

54. Hyvärinen, J.; Shelepin, Y. Distribution of visual and somatic functions in the parietal associative area 7 of the monkey. *Brain Res.* **1979**, *169*, 561–564.
55. Motter, B.C.; Mountcastle, V.B. The functional properties of the light-sensitive neurons of the posterior parietal cortex studied in waking monkeys: Foveal sparing and opponent vector organization. *J. Neurosci.* **1981**, *1*, 3–26.
56. Sakata, H.; Shibutani, H.; Kawano, K. Functional properties of visual tracking neurons in posterior parietal association cortex of the monkey. *J. Neurophysiol.* **1983**, *49*, 1364–1380.
57. Baizer, J.S.; Desimone, R.; Ungerleider, L.G. Comparison of subcortical connections of inferior temporal and posterior parietal cortex in monkeys. *Vis. Neurosci.* **1993**, *10*, 59–72.
58. Komatsu, H.; Wurtz, R.H. Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *J. Neurophysiol.* **1988**, *60*, 580–603.
59. Newsome, W.T.; Wurtz, R.H.; Komatsu, H. Relation of cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. *J. Neurophysiol.* **1988**, *60*, 604–620.
60. Goodale, M.A.; Milner, A.D. Separate visual pathways for perception and action. *Trends Neurosci.* **1992**, *15*, 20–25.
61. Tootell, R.B.H.; Reppas, J.B.; Kwong, K.K.; Malach, R.; Born, R.T.; Brady, T.J.; Rosen, B.R.; Belliveau, J.W. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J. Neurosci.* **1995**, *15*, 3215–3230.
62. Paradis, A.L.; Cornilleau-Pérès, V.; Droulez, J.; van de Moortele, P.F.; Lobel, E.; Berthoz, A.; le Bihan, D.; Poline, J.B. Visual perception of motion and 3-D structure from motion: An fMRI study. *Cereb. Cortex* **2011**, *10*, 772–783.
63. Schenk, T.; McIntosh, R.D. Do we have independent visual streams for perception and action? *Cogn. Neurosci.* **2010**, *1*, 52–62.
64. Watson, J.D.G.; Myers, R.; Frackowiak, R.S.J.; Hajnal, J.V.; Woods, R.P.; Mazziotta, J.C.; Shipp, S.; Zeki, S. Area V5 of the human brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cereb. Cortex* **1993**, *3*, 79–94.
65. Plant, G.T.; Nakayama, K. The characteristics of residual motion perception in the hemifield contralateral to lateral occipital lesions in humans. *Brain* **1993**, *116*, 1337–1353.
66. Federspiel, A.; Volpe, V.; Horn, H.; Dierks, T.; Franck, A.; Vannini, P.; Wahlund, L.-O.; Galderisi, S.; Maj, M. Motion standstill leads to activation of inferior parietal lobe. *Hum. Brain Mapp.* **2006**, *27*, 340–349.
67. Kaas, J.S. Human visual cortex: Progress and puzzles. *Curr. Biol.* **2004**, *5*, 1126–1128.
68. Culham, J.C.; Valyear, K.F. Human parietal cortex in action. *Curr. Opin. Neurobiol.* **2006**, *16*, 205–212.
69. Farrer, C.; Frey, S.H.; van Horn, J.D.; Tunik, E.; Turk, D.; Inati, S.; Grafton, S.T. The angular gyrus computes action awareness representations. *Cereb. Cortex* **2008**, *18*, 254–261.
70. Bonda, E.; Petrides, M.; Ostry, D.; Evans, E. Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *J. Neurosci.* **1996**, *16*, 3737–3744.
71. Grossman, E.; Donnelly, M.; Price, R.; Pickens, D.; Morgan, V.; Neighbor, G.; Blake, R. Brain areas involved in perception of biological motion. *J. Cogn. Neurosci.* **2000**, *12*, 711–720.

72. Grèzes, J.; Fonlupt, P.; Bertenthal, B.; Delon-Martin, C.; Segebarth, C.; Decety, J. Does perception of biological motion rely on specific brain regions? *NeuroImage* **2001**, *13*, 775–785.
73. Pelphrey, K.A.; Mitchell, T.V.; McKeown, M.J.; Goldstein, J.; Allison, T.; McCarthy, G. Brain activity evoked by perception of human walking: Controlling for meaningful coherent motion. *J. Neurosci.* **2003**, *23*, 6819–6825.
74. Johnson, M.H. Biological motion: A perceptual life detector? *Curr. Biol.* **2006**, *16*, R376–R377.
75. Senior, C.; Barnes, J.; Giampietro, V.; Simmons, A.; Bullmore, E.T.; Brammer, M.; David, A.S. The functional neuroanatomy of implicit-motion perception or representational momentum. *Curr. Biol.* **2000**, *10*, 16–22.
76. Kourtzi, Z.; Kanwisher, N. Activation in human MT/MST by static images with implied motion. *J. Cogn. Neurosci.* **2000**, *12*, 48–55.
77. David, A.S.; Senior, C. Implicit motion and the brain. *Trends Cogn. Sci.* **2000**, *4*, 293–295.
78. Kourtzi, Z.; Kanwisher, N. Implied motion activates extrastriate motion-processing areas: Response to David and Senior (2000). *Trends Cogn. Sci.* **2000**, *4*, 295–296.
79. Senior, C.; Ward, J.; David, A.S. Representational momentum and the brain: An investigation into the functional necessity of V5/MT. *Vis. Cogn.* **2002**, *9*, 81–92.
80. Dobrez, L.; Dobrez, P. Canonical form and the identification of rock art figures. *Am. Indian Rock Art* **2013**, *39*, 115–129.
81. Zeki, S. *Inner Vision: An Exploration of Art and the Brain*; Oxford University Press: Oxford, UK, 1999.
82. O’Craven, K.; Rosen, B.R.; Kwong, K.K.; Treisman, A.; Savoy, R.L. Voluntary attention modulates fMRI activity in human MT-MST. *Neuron* **1997**, *18*, 591–598.
83. Jarrett, C.B.; Senior, C. Implicit motion perception in schizotypy and schizophrenia: A representational momentum study. *Cogn. Neuropsychiatry* **2002**, *7*, 1–14.
84. Rao, H.; Han, S.; Xue, Y.; Gu, H.; Cui, Y.; Gao, D. Engagement of the prefrontal cortex in representational momentum: An fMRI study. *NeuroImage* **2004**, *23*, 98–103.
85. Krekelberg, B.; Vatakis, A.; Kourtzi, Z. Implied motion from form in the human visual cortex. *J. Neurophysiol.* **2005**, *94*, 4373–4386.
86. Schiff, W.; Caviness, J.A.; Gibson, J.J. Persistent fear responses in rhesus monkeys to the optical stimulus of “looming”. *Science* **1962**, *136*, 982–983.
87. Kelly, M.H.; Freyd, J.J. Explorations of representational momentum. *Cogn. Psychol.* **1987**, *19*, 369–401.
88. Tyll, S.; Bonath, B.; Schoenfeld, M.A.; Heinze, H.-H.; Ohl, F.W.; Noesselt, T. Neural basis of multisensory looming signals. *NeuroImage* **2013**, *65*, 13–22.
89. Rind, F.C.; Simmons, P.J. Seeing what is coming: Building collision-sensitive neurones. *Trends Neurosci.* **1999**, *22*, 215–220.
90. Duffy, C.J.; Wurtz, R.H. Sensitivity of MST neurons to optic flow stimuli. I. Continuum of response selectivity to large-field stimuli; II. Mechanisms of response selectivity revealed by small-field stimuli. *J. Neurophysiol.* **1991**, *65*, 1329–1359.
91. Anderson, K.C.; Siegel, R.M. Optic flow selectivity in the anterior superior temporal polysensory area, STPa, of the behaving monkey. *J. Neurosci.* **1999**, *19*, 2681–2692.

92. Maier, J.X.; Neuhoff, J.G.; Logothetis, N.K.; Ghazanfar, A.A. Multisensory integration of looming signals by rhesus monkeys. *Neuron* **2004**, *43*, 177–181.
93. Dobrez, L.; Dobrez, P. Canonical figures and the recognition of animals in art and life. Forthcoming in *Boletín del Museo Chileno de Arte Precolombino*.
94. Dobrez, L. The Reception of Visual Representations. In *XXII International Valcamonica Symposium*; Edizioni del Centro: Capo di Ponte, Italy, 2007; pp. 115–121.
95. Dobrez, L. A Rock Art Typology: Narrative and Non-Narrative Figural Representation. In *Aesthetics and Rock Art III Symposium*; Heyd, T., Clegg, J., Eds.; Ashgate Publishing: Hampshire, UK, 2008; pp. 29–33.
96. Dobrez, L. “Good to Think” Representational Modalities. In *Global Rock Art FumdhaMentos IX*; Fundação Museu do Homem Americano: São Raimundo Nonato, Brazil, 2010; pp. 284–294.
97. Dobrez, L. Looking at our looking: A new approach to the definition of a rock art scene. *Am. Indian Rock Art* **2011**, *37*, 251–264.
98. Dobrez, L. Towards a More Rigorous Definition of Terms: Are There Scenes in European Palaeolithic art? In *L’Art Pléistocène dans le monde*; Actes du Congrès IFRAO, Tarascon-sur-Ariège, septembre 2010. Numéro spécial de Préhistoire, Art et Sociétés, Clottes, J., Ed.; Bulletin de la Société Préhistorique Ariège-Pyrénées, LXV-LXVI, Ariège, France, 2010/2011; p. 316 (Abstract); pp. 1837–1851 (CD full text).
99. Keyser, J.D.; Dobrez, L.A.C.; Hann, D.; Kaiser, D.A. How is a picture a narrative? Interpreting different types of rock art. *Am. Indian Rock Art* **2013**, *39*, 83–99.
100. Keyser, J.D.; Kaiser, D.A.; Dobrez, L. Biographic rock art tallies: Explicit, implicit, or inferred narrative? Forthcoming in *American Indian Rock Art* 40, 2014.
101. Lenssen-Erz, T. Coherence—A constituent of “scenes” in rock art. The transformation of linguistic analytical models for the study of rock paintings in Namibia (with Commentaries and Reply). *Rock Art Res.* **1992**, *9*, 87–105.
102. *Hermeneutics: The Handwritten Manuscripts by Friedrich Schleiermacher*; Kimmerle, H., Ed.; Duke, J., Forstman, J., Translation; Scholars Press: Atlanta, GA, USA, 1977.
103. Gadamer, H.-G. *Truth and Method*, Second Revised Edition.; Weinsheimer, J., Marshall, D.G., Translation; Continuum: New York, NY, USA, 1993.
104. Lewis-Williams, J.D. Cognitive and optical illusions in san rock art research. *Curr. Anthropol.* **1986**, *27*, 171–178.
105. Bednarik, R.G. *Rock Art Science: The Scientific Study of Palaeoart*; Aryan Books: New Delhi, India, 2007.
106. Helvenston, P.A. New narratives for rock art studies. *Rock Art Res.* **2013**, *30*, 161–163.
107. Helvenston, P.A. Differences Between Oral and Literate Cultures: What We Can Know About Upper Palaeolithic Minds. In *The Psychology of Human Behavior*; Bednarik, R.G., Ed.; Nova Publishers: New York, NY, USA, 2013.
108. Ong, W.J. *Orality and Literacy: The Technologizing of the Word*; Routledge: London, UK, 1982.
109. Donald, M. *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*; Harvard University Press: Cambridge, MA, USA, 1991.

110. Ramachandran, V.S. Mirror Neurons and Imitation Learning as a Driving Force Behind “The Great Leap Forward” in Human Evolution. Available online: **Edge** 69, June 1, 2000 (<http://edge.org/documents/archive/edge69.html>) (accessed on 4 December 2013).
111. Bednarik, R.G. The earliest evidence of palaeoart. *Rock Art Res.* **2003**, *20*, 89–135.
112. Bednarik, R.G. Megafauna depictions in australian rock art. *Rock Art Res.* **2013**, *30*, 197–215.
113. Leroi-Gourhan, A. *The Dawn of European Art: An Introduction to Palaeolithic Cave Painting*; Champion, S., Translation; Cambridge University Press: Cambridge, UK, 1982.
114. Hagen, M.A. *Varieties of Realism: Geometries of Representational Art*; Cambridge University Press: Cambridge, UK, 1986.
115. Köhler, W. *Gestalt Psychology: An Introduction to New Concepts in Modern Psychology*; Liveright: New York, NY, USA, 1947.
116. Michotte, A. *The Perception of Causality*; Miles, T.R., Miles, E., Translation; Methuen: London, UK, 1963.
117. Thinès, G.; Costall, A.; Butterworth, G. *Michotte’s Experimental Phenomenology of Perception*; Lawrence Erlbaum Associates: New Jersey, USA, 1991.
118. Hubbard, T.L.; Blessum, J.A.; Ruppel, S.E. Representational momentum and Michotte’s (1946/1963) “launching effect” paradigm. *J. Exp. Psychol.* **2001**, *27*, 294–301.
119. Peeters, R.; Simone, L.; Nelissen, K.; Fabbri-Destro, M.; Vanduffel, W.; Rizzolatti, G.; Orban, G.A. The representation of tool use in humans and monkeys: Common and uniquely human features. *J. Neurosci.* **2009**, *29*, 11523–11539.
120. Keyser, J.D.; Klassen, M.A. *Plains Indian Rock Art*; University of Washington Press: Seattle, WA, USA, 2001.
121. Dobrez, L. Framing and the Organization of Space in Representation. In *Global Rock Art Fumdhamentos IX*; Fundação Museu do Homem Americano: São Raimundo Nonato, Brazil, 2010; pp. 651–665.
122. Ramachandran, V.S.; Hirstein, W. The science of art: A neurological theory of aesthetic experience. *J. Conscious. Stud.* **1999**, *6*, 15–51.
123. Ramachandran, V.S. *The Emerging Mind: The Reith Lectures 2003*; Profile Books: London, UK, 2003.
124. Onians, J. *Neuroarthistory: From Aristotle and Pliny to Baxandall and Zeki*; Yale University Press: New Haven, CT, USA, 2007.
125. Blackburn, T. Biopsychological aspects of chumash rock Art. *J. Calif. Anthropol.* **1977**, *4*, 88–94.
126. Hedges, K. Phosphenes in the context of native american rock art. *Am. Indian Rock Art* **1982**, *7/8*, 1–10.
127. Bednarik, R.G. On the nature of Psychograms. *Artefact* **1984**, *8*, 27–31.
128. Lewis-Williams, J.D.; Dowson, T.A. The signs of all times: Entoptic phenomena in upper palaeolithic art (with commentaries and reply). *Curr. Anthropol.* **1988**, *29*, 201–245.
129. Hodgson, D. Art, perception and information processing: An evolutionary perspective (with commentaries and reply). *Rock Art Res.* **2000**, *17*, 3–34.
130. Watson, B. *Universal Visions: Neuroscience and Recurrent Characteristics of World Palaeoart*. Unpublished Ph.D. Dissertation, University of Melbourne, Melbourne, Australia, 2009.

131. Summers, D. *Real Spaces: World Art History and the Rise of Western Modernism*; Phaidon: London, UK, 2003.

© 2013 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).