VISUAL EFFECTS AND PHENOMENOLOGY
OF PERCEPTUAL CONTROL

Jay Lampert (University of Guelph and Duquesne University)

INTRODUCTION

I am interested in phenomenology of perception for digital visual effects programs like Adobe After Effects, and editing programs like Avid Compositor, and Final Cut. There is discussion in philosophy and film of the aesthetics of digital effects, but little on the perception of compositors.[1] The best resource for phenomenologists is to try out the programs, read training manuals, online tutorials[2] and trade magazines like *Cinefex*, or to take a few Introductory film school courses in these programs.[3]

There are many models of perception, in the sense that various parts of the process of perception have structures, some of which might be generalizable. Neurophysiology, information processing, perspectivism in painting — each offers categories for a theory of perception. These are only parts of perception, so they are models, not descriptions of perception. My question here is how visual effects editing provides a distinctive model of perception.

One could ask how the viewer perceives visual effects, or about visual effects as intentional objects. Instead, I will ask how compositors who make visual effects control their perceptions. This paper has three sections: the art of perceptual control; the After Effects model of perception, namely perception by layers and transparency; and implications for time and perspective.

THE ART OF PERCEPTION CONTROL

We are the cause of our perceptions in many ways. We draw pictures and then perceive them. We take a few steps and see things we did not see before. We concentrate, squint, drink beer, and hum. We colorize old movies. If digital effects were just a way of looking for new images or altering old ones, images which we then perceive with our sense-organs in the

usual way, then it would tell us little that is substantially new about perception. However, there are prima facie reasons to think that digital editing might constitute a new kind of perceptual control.

First, if we believe Malabou's hypothesis in *What Should We Do with Our Brain*,[4] then if we monitor and manipulate our brains while we perceive, we might end up with a different consciousness of perception, as well as a different set of perceptions. Digital effects editing might be like that. And if we believe Andy Clark in *Supersizing the Mind*,[5] a person's perception includes not only what is in her brain (or mind), but also includes perceptual resources like libraries and other people. By analogy, the work an editor does at the editing suite might count not just as cognitive pre-production, after which perception would take place, but as perception itself.

Second, non-linear digital editing models real perception more closely than the old linear editing of filmstrip did. It is faster, so it generates perceptions in realtime, yet because it never alters the original footage by cutting it physically (that is the sense in which it is non-linear), it preserves past images in a way that neither linear film editing, nor human memory, can do. In short, there is reason to think that non-linear editing genuinely models real perception, but with new features. Insofar as an editor not only offers perceptions to a subsequent audience, but also controls her own primary perception in realtime, there is something novel for phenomenology.

Perceptual compositing and image-machines suggest Deleuzian assemblages[6]: composited images always have one more level to add, and one to subtract. However, Deleuze's own analyses of cinema do not cover two of the key elements in compositing.

First, as has often been remarked, Deleuze rarely discusses the technological side of filmmaking. He does remark on events like the introduction of inexpensive video cameras into 1970's film culture. Nevertheless, his whole approach to cinema is founded on the argument, contra Bergson, that one central element of technology is irrelevant to the nature of cinema: the fact that film is projected in 24 interrupted frames/second, none of which moves, is for Deleuze irrelevant to whether film images "move." While many features of Deleuze's treatments of cinema presuppose technology in various ways, his concerns are more about viewing and thinking cinema than making it.

Second, while Deleuze's research is guided around the two types of cinematic image — movement and time-images — there is an ambiguity over what Deleuze means by "image."

An image is certainly not limited to what we sometimes call the composition of a frame, namely the spatial arrangement of figures inside a 4:3 rectangle. But is an image in general more like framing or more like a cut? Is an image more like a percept, an affect, or a thought? It is hard to say whether the overlap of image-layers is more like any-layer-whatever depth of field, or more like a grain-of-the-voice semiotic, or more like a history-of-effects hermeneutic.

Most important, once Deleuze says that an image is a function of sheets of time, and mounts past and future, i.e. non-present times, simultaneous with the present, an image cannot be the sort of thing we perceive at a given time. And if an image, for Deleuze, is thus not present, to any one at any time, then images need a kind of phenomenology outside the region of perceptual fields, passive syntheses, foreground figuration, and all the other categories that go with the action-image. If the idea of an "image" in the time-image is less about the content of awareness and more about the repetition of de-synchronizable collectives, and less about perceptual givens than perceptual freedoms, then the paradigm image is not on screen at the moment, but something like the time-consuming teamwork of layer-building.

Now, if editors did not create primary perception, but only tinkered with natural images in media received from external sources, then Bazin and Cavell would still be right: film would be a realist medium. Of course, lenses, lights, framing, and animation shape recorded images, and editing alters them, but whatever gets recorded would really have been there in front of the lens, from the perspective where the cinematographer's eye was, and where the audience will take up the point of view. However, it is not always the case that the editor sees the images first by natural perception and subsequently manipulates them. For example, an editor can manipulate images by algorithms without seeing them first. Indeed, when an effect is computationally expensive (some individual frames of visual effects in *Iron Man 3* (2013)[7] took over ten hours for the computer to render), the editor will likely previsualize the effect image in shortcut "preview" form;[8] on her workstation, she perceives only a sketch of the effect (more blurred or jittery than the polished render). She might in preview mode increase the exposure to see it more clearly — known as "slamming the comp." This high exposure image is part of the decision workflow but will not appear in the final perception, which exists only later on. In short, digital editing does not just manipulate natural images, but also builds original perceptions.

In fact, controlling perceptual syntheses is neutral as to whether perceptual objects are real, since synthesis constitutes both real and unreal intentional objects. Digital cinema is in

principle neither more nor less realistic than celluloid. To take one complex example, photo-realistic computer animation of human faces is now possible for still images, but too expensive for moving images. The problem in assessing degrees of realism raises what robotics calls the "Valley of the Uncanny": as images of faces become more realistic, they look better, until they get very close to fully real, at which point they look creepy, unless they get perfect. To make facial movements look perfect, some animators think,[9] would require artificial intelligence software, to simulate subtle preconscious cognitive micro-gestures on an artificial face. Of course, the issue of what looks realistic on screen is different from what is real. Obviously, an animator knows his creatures are not real; but after effects moves, like re-lighting elements or layering visual fields, can equally be in the service of the real, or the unreal, or some hybrid. Indeed, the ontological vocabulary in the digital effects industry takes some getting used to. "Reference" does not mean denoting real objects, but using pre-existing images to build CG versions, as someone animating Air Force One "references" Google images of it from many angles to build a 3-D animated version. To make an image look realistic is to "sell" it. Normally, images look realistic when they are dirty (in human perception, clean looks fake), so there are realism-generating programs that add smudges, cracks, and fumes. There are exceptions: "reference" shows that the real Air Force One is kept cleaner than all other airplanes, so animators cannot use the dirt-is-real trick to sell it.[10] In sum, when one composites perceptions for oneself, these have the same likelihood of realism as one's other perceptions.[11]

If it were possible to generalize from the phenomenology of compositors to the phenomenology of live perception, we might begin to want to control and edit our everyday perceptions. Nevertheless, the challenge in a topic like this is to drive the conclusions to the limit, yet without exaggerating. I want to analyze interesting features of effects work without the metaphor of a grand transformation of the human species. For example, green screen imaging reveals interesting features of background contingency, but there is no green screen inside the mind. There is no way to control all perception and still be in the world. However, digital editing does control some perception, and that by itself might without exaggeration be called the biggest event of the 21st Century so far. The idea is not just that the audience decides what is in the film (as when focus groups influence editors, or the way an Expo '67 audience voted on which pre-filmed ending they wanted to see), or that open-source footage

allows viewers to edit their own versions or sequels (the way Harry Potter fans make puppet shows on Youtube.[12]) The idea is to pin down the specific art of control in visual effects.

Analyzing user phenomenology is complicated, since programs like After Effects are designed for varied users: for Old School flatbed editors and mixers, the screen shows icons of dials; for math-averse artists, it shows paintbrush icons to click and drag over images; for programmers, numerical calculations; for Deleuzian assemblers, layers and strata.

Of course, there are limits to perceptual control. In the big picture, the CG sensorium offers less diversity than a carbon-based life form gets walking through a forest. And in those parameters where digital control can tune image detail more finely than eye control can, it can become too fine to be perceived at all. No doubt we are conscious of finer detail than we normally attend to, so there is room to benefit from artificial perceptual fine-tuning. Still, there is no value, for example, controlling sense-content beyond our peripheral vision. Roland Barthes discusses something analogous: not every tiny phonetic difference can make a difference in meaning. Barthes calls this the "security margin,"[13] or the "edge of the field of dispersal." By analogy, in digital editing programs, the view screen shows a "TV safe" border: before TV's were flat screen, their rounded sides would fail to show what was on the edges of the composition, so the editor would not put important content in that unsafe zone. Part of analyzing the control of perceptual parameters is to know where to stop.[14] The arguments of Dreyfus and Dennett still hold: the more we can do, the more we cannot do it. The fastest programmers cannot keep up with the speed of consciousness. It took a thousand programmers a year to make visual effects for a lousy movie like *Lord of the Rings*. But my point is not that consciousness can be replaced by controlled digital editing, just that editing can give consciousness a perceptual experience it could not have on its own, the way Vertov said the camera eye can see what the flesh eye cannot. A special effect-perception is almost a shimmering signifier in Barthes' sense,[15] and almost a saturated intuition in Marion's. Too perceptual for hermeneutics, too imagistic for *différance*, synthetic perception is a dialectic of also and insofar as, as Hegel says.

Béla Balász wrote in the 1940's[16] that cinema still needed a theoretical aesthetics even 50 years into its history — in part so audiences would know what is possible, and demand better movies. This is true today for controlled imagery. It is not enough either to be dazzled by special effects, or to reject them on principle.[17] Our youthful dream of smart-brain and virtual reality implants, cyborg extensions, and space travel with contacts never panned out. I

hope some artist implants T-cells into art-lovers' eyes for a new Op art. However, if digital editing programs are the only technological perceptual enhancements we have for now, we can still demand better perceptions. The fact that certain editing styles go out of fashion, like fades and wipes, reminds us that certain perceptions do too. One no longer glances, stares, or does doubletakes as people once did. It is a timeless norm not to blink or cut on the action, but some perceptions are diachronically emergent, and phenomenology needs to be on their cutting edge.

OBJECTIONS

There will no doubt be reservations about this idea of controlled perception, either on grounds that digital compositing adds nothing new to natural perception, or on grounds that what it adds is false. The former divides into two: that consciousness already does what compositing does, and that older art forms, from painting to analogue editing, already did it.

The first variation is to say metaphorically that consciousness already edits itself,[18] and that at least in imagination, we already vary images on parameters like colour, scale, and motion path, so the new technology producing perceptual variations is no big deal. Phenomenologically, though, it is not certain that we can imagine as many variations as we can perceive, like a hundred shades of green. Furthermore, as Husserl argues, even if we can imagine as much as we can perceive, imagination is not a subspecies of perception, but has different noetic properties. It is not obvious that an imagination content of emerald green looks emerald green in the same way that perception content looks emerald green. In digital editing, we do not imagine a hundred shades of green and decide which one to see; we dial through the shades and see them all, then pick one for the composition, and see it in the comp screen. The fact that After Effects controls perceptions, not imagination, is substantial.

To be sure, this happens too when painters choose among tubes, so even if consciousness does not already vary perception in the same way that effects editing does, perhaps we should admit that painting already does what effects editing does. However, in the view of computer animators, the difference between corporeal and computer painting is that computer painters go straight to making the image, whereas hand painters waste time first "smearing goo on a surface."[19] The point is even clearer if we take the paradigm of house

painting. On this paradigm too, the painter controls what we will perceive, but in this case, the painter first alters a pre-existing object and then perceives it, whereas the computer painter alters the perceptual content itself from the moment it first exists.

It is always difficult to assess the degree of difference between two phenomena (like hand painting and computer animation), just because it is always the case that dichotomies, like physical and virtual imaging, or like image-production and image-perception, or like active and passive, or control and receptivity, can be deconstructed. But that does not mean that all phenomena are ultimately the same, or that novelties are all really classical. Every art form starts with some sensuous material and turns it into something different in its own way. Differences on a continuum, or differences between technologies, or between user interfaces, are significant differences.

And yet it remains difficult to distinguish between digital image control and image control by a range of technologies including hand drawn animation, analogue compositing, or photochemical trick photography. Pre-CG movies like *Blade Runner* (1982), after all, required just as much multi-tasking visualization as digital compositing does. One of the creators of Photoshop, John Knoll, remembers physical editing as if it were already a precursor of digital multitasking:

> [Pre-digital] Optical compositing was always a performance. Load this element and this projector head on the [optical] printer, wind to this frame…, set the focus ever so slightly out to soften that, put this color filtration on here, shoot with this exposure, then wind back — and you'd do all of that for dozens of passes! And invariably there would be some mistake; so you'd fix that, and then something else would go wrong.[20]

Nevertheless, one difference between the more physical optical compositing and hand-drawn animation on the one hand, and the more virtual programmable effects on the other, concerns what the user has to know about the motion of objects. In hand-drawn animation, one can make an eye tear up as one wishes; to program the computer to make the eye tear up, one has to know, for example, how the meniscus layer of the eye socket creates liquid suction between the eyeball and the surrounding tissue, so the programmed keyframes will engage corneal controls, eyelid controls, and lid follow-through.[21] In programmed motion graphics, if one wants objects to undergo complex curving movements rotating around one

another, one has to know whether the objects are rotating around a fixed point on the screen or around each other, and where their "anchor points" are. If one misjudges which relative rotations one wants, the algorithms will take the program at its word, and make something unintended. But if one knows what movement one wants to see, the program will iterate it easily. In general, hand-drawn animation allows one to draw what one wants to see without knowing exactly the logic of rotation; but then one has the hard job of making each frame oneself.

In one sense, visual effects are in principle independent of computer animation, since one can put after effects on live action footage. In *Iron Man 3*, for example, Iron Man saves a dozen people who fall out of a plane, guiding them together to the ground.[22] Animating this scene might have worked for long shots, but the director wanted close ups. Green screen might have seemed an obvious choice to produce the scene (one films actors against a green background, then removes the green, and replaces it with a CG environment). But it is difficult to do green-screen on people whose hair is blowing in the wind, since the green light of the background bends around and gets reflected in the hairs, so when it comes time to "key" or remove the green, green-removal leads to hair-removal. As an alternative, a typical stunt would be to wire actors to a wind tunnel; but wind-tunnels limit camera movement. So instead of these options, they filmed the Red Bull Skydiving team actually jumping out of planes (digitally painting out their parachutes later). The difficulty was that one jump yields only 45 seconds of film, so they filmed 60 jumps over 18 days. But then the weather and light reflections had to be consistent, so the live-action film had to be modified with effects: effects, but not full CG. Admittedly the line between computer graphics and effects editing is a fine one, since for this scene some depth cues did have to be animated.[23] The line is blurred even more in the technology of performance capture, where live actors' facial expressions are marked, digitized and transferred to computer, after which animators paint images over top of the geometry of the marks. This blurring of distinctions has created a problem for the Oscar category for "Animation."[24] And all this applies not only to big-budget special effects, but also to barely noticeable perceptual enhancements, when digital effects are used to add just a touch more twinkle to a live actor's eye, or a touch more reflection in the metal on a car door.[25]

Because the phenomenological lines are blended and the technologies are interactive, much of visual effects work consists in troubleshooting unintended consequences of interacting perceptual parameters.[26] For example, moving an object has implications for blurring,

though control of movement and blurring may require two keyboard operations.[27] Edgar Burcksen, who runs *CinemaEditor Quarterly*, says: "When I flop a shot to fix an eye line problem, it can compromise the light source the director of photography has established. When I blow up, reframe or zoom in on a shot, I'm messing with the depth of field. When I slow down or speed up a shot, I'm screwing with motion blur."[28] Parameter control can be a reveal, as they say in mystery movies, for phenomenology. Just as a close-up may create problems for a scene — it may, for example, create a tempo conflict (since listening to a long speech is reasonable, but watching the speaker's face for a long time is not), or it may create a lighting conflict (since lighting an actor for glamour often conflicts with the diegetic light) — in the same way, a layered-on visual effect may sit uneasily in its composition. In consequence, the distinction between independent pieces and non-independent aspects may not be fundamental ontology, as Aristotle and Husserl thought, but a dial to turn up or down. Controllable perception needs new categories of wholes and parts.

The second objection is that computer programs for digital effects do make a new technology, but tell us nothing about conscious perception. Effects technology might be criticized for technocentrism, in wanting to direct every detail.[29] It is true that there used to be hundreds of uncontrolled improvisations by crew members, from focus pullers to lighting grips, which can now be micro-controlled in post-production. But control is not necessarily opposed to chance or excess. In music, for example, the movements between, on the one hand, total serialism and algorithmic assists in composition, and on the other, aleatory composition and improvisation, no longer seem so exclusive. Like visual effects compositing, the point of both is to de-naturalize the balance of control and freedom. When it is said that these visual effects programs are "deep," it means that the user can go into any parameter and make decisions. The program does do some things automatically: if the animator tells it to smooth out a jerky camera movement, it will, without telling her, introduce a little motion blur, so as not to leave gaps as it alters perspective. But one can turn off any program heuristic. For every element, there is a dialogue box between consciousness and the program that controls its perception.

Still, it is true that perception control applies not to the film watcher, but to the compositing technician, and in that sense it might not apply to all forms of consciousness. Compositor phenomenology takes to the limit Kant's thesis that we get out of perception what we put into it. Of course, as the technology gets cheaper and invades all our electronic devices, and

we come to edit our perceptions through the windshield of the car and into the back of the fridge, we might all composite images with the same programs. But so far, digital compositors perceive differently than others. If we look at controlled perception in Adobe After Effects as a project file, we can see the edits on screen, examine the layers on the timeline panel, and see all the marks of the effects. We might even think of the project file as a work of art independent of the movie scene that renders it, just as Balász argued that a film script is a work of art independent of the film, or in the way that musical scores can be appreciated independent of performances (249). Once we turn the project file into a movie, we see the image as a completed perception, as a ready-built artifact without its layered sediments, and then, paradoxically, throwing away the ladder, it appears as a raw, immediate perception. The movie watcher sees, but the compositor alone looks. For that matter, there is often no one screen where the whole image exists until the last minute. *Iron Man 3* employed seventeen different effects studios, who could not always wait for the others' results, so three different effects vendors animated Iron Man's suit in various stages of destruction. Just as there is no one maker of a film even when the director is aptly called its "author," so perception control is a team effort.

One way to isolate the compositor's experience is to compare different programs: in Adobe After Effects, the compositor layers clips of footage over each other and sees the blend. In contrast, programs like Nuke use nodes[30]: the compositor sees on screen a flowchart of boxes representing clips of footage linked by the names of visual effects. She sees the image of the final mix, but does not generally see the layers show through each other one at a time.[31] The results can be the same across different programs, but the building process is different, which means the perception of perception-control comes during the control, not after it has been controlled. This is somewhat true of natural perception too, and also true of reason, decision and other noeses. (Perception itself may be a kind of decision.[32]) The experience is in the construction, not just the conclusion, and most of perception takes place before we are aware of the pictorial representation. One has to do math to know what it is, and so one has to composite images in order to evolve perceptually. Most viewers cannot tell whether an effects scene is based on miniature photography (as in *Skyfall*, 2012) or motion capture (*Iron Man 3*) or graphic animation (the goblins in *The Hobbit*, 2012). Look at a Bollywood dance scene; are the perfectly synchronized dancers in the back rows live performers, or digitally mass-produced by a program called Massive? The gap between compositors and movie

watchers is larger than the gap between painters and connoisseurs (just as the gap widens between tweeters, whose capacity for connected thought withers, and programmers of tweeting software, who hone those skills). It is not unreasonable that Making-of movies are often more interesting than the movies themselves, and philosophers may share more with the how-did-they-do-it geeks than with film critics.

THE AFTER EFFECTS MODEL:
LAYERING AND TRANSPARENCY

There are many digital operations in cinema: animation, rotoscoping (cutting objects out of the background and moving them around), colour correction and image distortion, not to mention digitally controlled camera movements, and data recording. I focus on just one compositing operation: the stacking of layers of image-materials. This model may turn out to be decade-specific. We may have to theorize a new art form every few years. Perhaps the novelty of layering art is not as radical as the change from print to silent movies,[33] i.e., from print to picture, since it controls sense-organs but does not switch to a new one.[34] Nevertheless, the category of layering has the potential to add new and essential elements to phenomenological description in general.[35]

The essence of the visual effects model of perception is that we manipulate visual material by letting one layer show through another. If theatre presents the whole stage, and film cuts the stage by angles,[36] layer art blends sheets. "Blending modes" produce overlapping colours, lightings, and motions. One can take footage, copy it, and blend it with itself in layers; if one blurs one layer and sharpens the other, one gets dreamy effects. If one lets an underlying layer of light show through, one can re-light a scene differently from the way it had been shot. In *Iron Man 3*, the character whom the effects team called "Volcano Man" was composited in layers of bones, nerves, muscles, veins, and skin; bright lights were built into the lower-down muscle layers, and the upper skin layers had degrees of transparency, so the hot light would burst out from inside his body. Animators do study natural light, like the way it penetrates under human skin and scatters around the blood before it is reflected. But there is a difference. Hitchcock physically placed a lightbulb inside the milk in a glass in *Suspicion* (1941), to make the glass glow suspiciously of poison, but he could not have placed a

lightbulb inside Volcano Man's stomach. A traditional director of photography treats light as natural perception; he places a physical lightsource at a certain angle and distance from an object, then compensates for pools of shadow by adding studio lights, or enhancing natural sources like sunlight, or practicals like table lamps, compensating again for degrees of warmth with blue or orange filters, gauze and gobos. The visual effects compositor, in contrast, makes light and shadow by stacking bright and dark layers with pools of transparency to allow, or block, blending. Studio lighting takes place largely from the outside, whereas layering light takes place in the inside, and this for phenomenology is crucial.

Layering and transparency are effectively the same concept, since layering is not possible without the transparency of the layer on top. Phenomenologically, we see things only when the things between us and them are transparent. We see Coke through a glass, the glass through water droplets on its surface, the glass and droplets through our eyelashes. Normally, transparency is by definition unnoticed. But in perceptual control, we can either make an opaque object (or the shape around it: the mask) transparent by assigning it to the Alpha channel rather than to RBG (red-blue-green) channels, or make a transparent shape opaque. We decide what degree of transparency to make the shape, and what degree of the layer underneath will show through. And then we control the transparency of that underlying layer, so the layer underneath it can show through in turn. Many features of perception (colour, brightness, sharpness, graininess, etc.) can be manipulated by doing nothing more than manipulating transparency.[37]

Generally, perception by layers is determined by which layer is on top, due to the render order of the program. A layer closer to the surface may undo an effect that was on a lower layer, so one may have to make a group of layers into a "Nested Comp," or "Parent" certain layers to others. Or one can parent a nest of compositions to a "Null object" or an invisible layer, a hierarchy without a parent, so the synthesizing element will be empty and not get in the way. One example is the effect called Find Edges, which accentuates edges on an object; it will not find the edge of an object that has replaced a different object on a lower layer. If one wants Find Edges to find its edge, one has to render the effect on the lower object before it was replaced.

At the onset of cinema, some filmmakers pushed for uniquely cinematographic ideas, not just to use cinema as photographed theater. It did not matter whether cinema had first been designed for scientific analysis, military surveillance, or diverting the proletariat: it de-

served novel aesthetic treatment. I do not think there should be a rule against using one art form merely as a prop for another, and not all book-to-film adaptations are wrongheaded. But even if it is not the only legitimate goal, it is clearly worth experimenting with uniquely cinematographic features to see what they can yield, and so by analogy it is worth trying to design layers and blending modes for their distinctive properties, perhaps to reveal characters and dramas, or clues and backstories, concealed and revealed in the layers. Instead of merely adapting traditional film stories for layer-art, it would be interesting to see what happens both to perception and to narrative if we were to design some films specifically for layers, in the way certain films were designed specifically for split screen, or Technicolor.[38] When Cinemascope was introduced, for example, it had the effect that when the camera panned quickly across a wide landscape, the background was blurry. One response was to mitigate the effect and conceal the artifact; some directors decided to pan more slowly, or to put nothing important in the background, to avoid the blurred background effect. In contrast, Douglas Sirk's alternative was to have people in the background move more quickly, accentuating the blur,[39] making the side-effect into its own aesthetic idea.

Up to this point, I think, layering itself has not become an aesthetic idea. Superimposing images, of course, goes back to the early films of Meliès. But layering image-elements is not the same as superimposing whole images. What will a layered image, qua layered, look like, so as not to pander to existing eye-usage?[40] Could we layer extreme foregrounds against extreme backgrounds without depth continuity in the middle, to force the eye to provide its own mid-tones,[41] or live without the middle, to see the layered prose of the world? We could, of course, thematize the fact of layering by disentangling the layers in a perceptual field and showing them in succession, or on split-screen, but that would show *that* there are layers, without actually layering them. The challenge is to make the layers visible, without separating them in a non-layered presentation. The current eye wants one or the other: invisible editing, or separated presentation; the new eye needs to see a difference without that difference.[42] One paradigm is Godard's *Histoire(s) du Cinéma* (beginning in 1988), though it was made with videotape editing rather than digital layering. Where normal documentaries about cinema show old favorites and lesser known movie scenes, Godard piles scenes on top of other scenes so we can barely make them out. He darkens irises around the scenes so we cannot make them out at all, or raises the contrast of light and dark so we cannot tell if there

is even a scene on screen — all the while, his voice-over appeals to the fatal beauty of the image. It is infuriating, an acquired taste, and an example of the art of layering.

This is an urgent question, because not many frames of the average film today come straight out of the camera, whether we look at Hollywood movies, or works by video artists like Jennifer Steinkamp, Jeremy Blake, and Takeshi Murata.[43] The visual effects programs are no longer expensive, and Photoshop has already made it as natural to control images as to capture them.[44] Cameras themselves are now designed for in-camera and subsequent image-control. The visual effects post-production team does pre-visualization during pre-production that shapes what the director shoots live in production. No doubt something is lost with any new technology. When silent films gave way to sound, dialogue scenes became static by comparison with action scenes in silent movies; and to prevent the loud noise of strong lights, high contrast expressionistic lighting of silent films gave way to a softer "glamour" aesthetic. But just as film itself added a new region of phenomenology, namely the camera eye, and talkies introduced the microphone ear, at some point layering art will extend phenomenology into the editor eye.

TWO PHENOMENOLOGICAL CATEGORY SHIFTS

Time[45]

Visual effects on space are obvious. William Brown's book *Supercinema* shows, for example, how digital cinema no longer needs cuts. The traditional 700-foot reel of 35 mm film only lasted ten minutes, after which there had to be a cut. Digital hard drives can record a feature length film without cuts, but even a digital camera, being physical, cannot pass through walls without a cut. Computer graphics, in contrast, can take an image continuously through walls. Brown suggests that movies today have cuts only to make older audiences feel comfortable.[46]

Following Deleuze, cinema should control not just action in space, and not just action over time, but time as such. Filmmakers have always controlled temporal density, i.e. frames per second,[47] and have added or removed frames to stretch time, for example to synchronize dancers. But effects-makers can now use a function called "Timewarp" to "varispeed" footage. The zombies in *World War Z* (2013), for example, were "re-timed," so that after the actors

(modern dancers were hired to portray the zombies closest to the camera) were recorded, one eye or limb was digitally delayed relative to the other, to enhance zombie-likeness.[48]

Compositors also control a graphic representing only the temporal features of a moving image. For example, if an object moves for ten seconds along a spatial path at a fixed rate of speed, the graphic time-line will appear as a straight diagonal. One can put the cursor over the line and drag and twist the curve — not twist the motion path, but the time-path.[49] This changes the patterns of acceleration and deceleration, which the time-curve now forces to fit the ten seconds. The effects compositor does not just stretch the painting over time, she paints with time, the way traditional Directors of Photography say they "paint with light."[50]

As always, technical difficulties arise, and each one points to an overlooked aspect of time. For example, an optical illusion arises when an object moves along a jagged path: when the object comes to a point on the jag and bounces back out, it appears to accelerate, even though its speed is constant. Heuristically, this is partly because in the physical world, we slow down when we approach a wall, and accelerate when we bounce. We expect to see deceleration at the point, and when we do not, it seems too fast. The effects compositor compensates by using a feature called Easy In/Easy Out to lower the object's speed at the point. We can control the entry and bounce-back ourselves, or we can let the program do it with a feature called Easy Ease. But this creates a further problem; if the clip is still ten seconds long, and the object slows at the jag points, then it will move faster between the jags to get to the end at the right time. Trouble-shooting is inherent to the stages of computer-assisted perception: control, meta-control, and tinkering with unintended consequences. Sometimes there are too many to compensate for. If there are too many jags, the object may stutter, or the speed between jags could be too fast for the audio track. In such cases, we need to rethink what we want: to omit some jags, or change the clip length. It is when we work with it plastically that we discover attributes of time, like Easy Time.

To me, the most interesting function is time-blending to control motion blur, which amounts to layering different times simultaneously over time. Blur is a natural effect of camera recording; a fast object will change position while a single frame is being exposed: hence blur. But sometimes the editor wants more blur than the camera gave (e.g., if the camera had to be set at very fast shutter speed while the actor was running in front of a green screen, to prevent the green from seeping into his hair, then normal blur will be absent). The way to digitally add blur is by "calculating intermediate positions between frames, then blending

together these multiple copies of the layer."[51] Or as an alternative to extrapolating intermediaries, the program can take any number of frames before and/or after a given frame and blend these past and futures into the present. This way of smearing time is called "Echo Time," or to call it by my favorite brand name, "Wide Time." Manipulating simultaneity and succession reveals that time has the property of width. There are many time-control plug-ins for sale, but time can run into real money.

Perspective

What is direction if we can make shadows extend from non-existent objects? What are dimensions if we can paste two-dimensional scenes onto 3-dimensional geography? What is perspective if a program called Vanishing Point can wrap an image around any point or group of points?

Layering frees perspective and depth of field from the focal plane. Normally, depth of field means that from the plane of greatest focus, there is a finite distance in front, and behind, where the object remains relatively focused. Devising ways of detaching depth of field from focal plane has always been a part of filmmaking. Take the famous scene in *Citizen Kane* (1941) of Susan's suicide attempt. Kane arrives from the back door, and walks forwards, towards Susan in bed; in front of Susan, in extreme close-up, is the glass of poison. Strangely, both background and foreground (Kane and the poison) are in sharp focus, but the middle ground, where Susan is, is out of focus. How is this optically possible? The answer is that this shot is not typical depth of field, made by light and lenses. It is an in-camera matte shot. Robert Carringer explains: "First, the foreground was lighted and focused, and shot with the background dark. Then, the foreground was darkened, the background lighted, the lens refocused, the film rewound in the camera, and the scene reshot," and double exposed.[52] This type of trick layering is made into the norm by digital compositing.

Perspectivism in 2-D is illusionistic anyways, as is the distinction between foreground and background, but the issue is complicated when visual effects blur the distinction between 2-D and 3-D imagery. There are what are called 2 and a half-D images, which rotate a 2-D object around three axes, so the object appears to move in front of and behind other objects, but if it rotates, it is revealed to be paper-thin (known as "postcards in space"). Full 3-D imagery allows independent rotation of objects on three axes: Euler rotation. (Lidar scans, using lasers and radar, capture the contours, so every angle is ready to plug into a 3-D envi-

ronment.) One can either use "Auto-Orient" to keep an object facing the viewer, or "Custom view" to preview perspectives of one's choice. Like every function, 3-D reveals perceptual problems while it adds perceptual content. For example, if a scene has been shot or animated in stereo for 3-D (the parallax view), then if a figure has to be painted out and replaced with something else, it has to be done twice, from the two eyelines, and the resulting images cannot simply be lined up by "corner pinning." Compositors call this the problem of the "second eye": there is not just one camera eye, but two. Once again, the compositor is forced to know this, while the rest of us are more or less cyclopic. There are some questions of empirical psychology that to my knowledge have not been studied, but I would like to know whether compositors tend not to see the 3-D images on their screens *as* 3-D, since they are too aware of how they have been built up by layers, or whether they are more like piano tuners who can still hear melodies (though actually, I don't know if piano tuners' ears have been studied either). If movie watchers want to look backwards towards the compositor's experience of layering under the 3-D hood, we can always watch a 3-D movie in 2-D. In any case, in physically embodied life, it is because objects exist in three dimensions that there are backgrounds; in After Effects, it is because there are layer-defined backgrounds that objects exist in 3-D.

If Lacan is right that objects gaze back at us, as Merleau-Ponty says of Cézanne's painted orange,[53] it should make a difference that objects now gaze back in 3-D. Of course, image 3-D is not the same as reality 3-D. Image 3-D has layers of surface, but no inside (unless the animator has built inside layers).[54] And the backside of an image 3-D object may be missing, but it is not exactly hidden. Image 3-D is in-between cubism on the one hand, where we see the backsides of objects at the same time as we see their fronts, taking up an inconsistent set of simultaneous perspectives, and kinaesthetics on the other, where we have to move our bodies before we see the backsides.[55] In image 3-D, we see front and back successively (unlike cubism), but simply in perception (without embodiment), so the backside is not inconsistent with the perspective we have on the front. The depth of the world is available (albeit successively) to a single consistent perspective. A 2-D painting of an orange, which gazes back by resisting our gaze, is to a 3-D orange, which gazes back by admitting our gaze; as the Levinasian distant other is to the Virtual Reality accessible other. No doubt, both extremes — distance and access — are false ideals. Still, 3-D perceptual control is one point on the side of accessibility without reserve.

In addition to 3-D, there are 4-D images, i.e. 3-D images interacting with time-scale. There are what producers call 5-D images, where space and time parameters are traded off with a cost-dimension. There are 6-D images, where the sixth dimension is the technology's shelf-life.

Compositing also allows for novel meta-perspectives. When footage is dragged into the composition panel, you can introduce a virtual camera into the composition, then see the image inside the panel from the perspective of the virtual camera that is also inside the panel. The scene can be made to see itself from the perspective of the eye inside the scene. The inner camera can be animated to pan, or follow a defined "point of interest" on a virtual dolly or crane, or simulate a handheld camera eye. There are settings for its virtual lenses, zoom, aperture, and any parameter of a natural camera, plus some. For example, the virtual camera can mix focal lengths in the same image, which a lens cannot. The footage, which the virtual camera re-shoots from within, may already have been shot by a physical camera. Whether we call this a repeating camera eye, or mind without eye, it is a distinctive sort of self-reference. It is not like Russell's paradox, a set that includes itself as a member of itself, or like a knower knowing itself; it neither shadows-off to a new perspective nor is it auto-poetic feedback; it is not like a *mise en abîme* reproducing the same image from within, nor a meta-language where signifiers refer to other signifier-signified relations; it is not like framing a photograph, or like filming a filmmaker filming a movie, nor is it like filming a film projected onto a screen using a videocamera in the audience. It is the opposite of an in-camera edit like the one in *Citizen Kane*. This is an in-edit camera. We build a perception, *then* we introduce conditions of possible experience, before we ever perceive the polished image on screen.

CONCLUDING QUESTIONS

Just as novice photographers walk around with their fingers framing the world in front of their eyes, compositing initiates on a walk may try to peer through the surface of things in the real world to the layers seeping from beneath. To say that compositing is a model of perception does not mean that reality is merely a simulacrum of special effects, but it does pose questions for natural perception. If different properties and elements lie on different layers,

what does consistency mean? What does *Erlebnis* mean, or the transcendental object = X? What do motion and rest, light and dark, live and artificial, mean? Layering blurs the distinction between foreground and background, between direct and indirect, between showing something itself and showing it in a mirror or by its shadow. And if Balász were right that violence ought not to be shown directly (actually, he is probably not right), then layering would also blur the distinction between good and bad taste.

The resources for a phenomenology of controlled perception are still largely wasted. When we watch a film with a lot of visual effects, like the newest *Star Trek*, do we pay attention to the actors in close-up, and half-see the CG background; or do we pay attention to the effects background while only half-seeing the characters, treating the diegetic background as our cognitive foreground? Probably, we most often see things the traditional way, with the characters receiving our attention,[56] which means effects are still in its infancy as art, used as a means to an end, in spite of complaints that movies today have too many glitzy effects and not enough plot or character. Obviously it is true also that films that are supposed to run on plot do not have enough of it, and waste time with visual effects, as Effects Supervisors are the first to insist. But this is not because they are built on effects — they are still built on plot, but badly. After After Effects becomes an art in its own right, we can expect to see things more clearly.

---

1. Lev Manovich's *Software Takes Command* (New York: Bloomsbury, 2013) contains a subsection titled "Layers, transparency, compositing" (277-282), specifically dealing with After Effects, but even better material on this topic is found in the subsection called "Inside Photoshop" (124-147). Manovich goes a step farther than my thesis that perception is open to graphic design, and argues that perception is commanded by software. Holly Willis's *New Digital Cinema* (London: Wallflower Press, 2005) makes a few comments about layering and After Effects. Relevant issues are discussed in William Brown's *Supercinema: Film-Philosophy for the Digital Age* (New York: Berghahn Books, 2013).

2. This paper is intended to be read in conjunction with viewing two online tutorials of about 10 minutes each: (1) "Sub Surface Skin" in After Effects, tutorial 140, videocopilot.net, and (2) "Blend and Bend Time" in After Effects, by Evan Stern, creativecow.net. If these sources are at some point pulled from the web, other online tutorials on the topics of layering and time blending will do as substitutes.

3. I would like to thank the wonderful instructors at the Pittsburgh Filmmakers School who inspired me: Dean Mougianis, Will Zavala, John Cantine and others. I would also like to thank the editor of this volume and the two anonymous referees for their suggestions.

4. Catherine Malabou, *What Should We Do with Our Brain?*, trans. Sebastian Rand (New York: Fordham University Press, 2008).

5. Andy Clark, *Supersizing the Mind* (Oxford: Oxford University Press, 2008).

6. Patricia Pisters, in *The Neuro-Image: A Deleuzian Film-Philosophy of Digital Screen Culture* (Stanford: Stanford University Press, 2012), interprets Deleuze's neuro-image as cinema that presents the inside of a brain, construing consciousness and interpretation as synaptic remixing and database re-ordering.

7. Jody Duncan, "Rough Around the Edges," *Cinefex* 134 (2013): 12-44.

8. Previews are essential to perceptual working-with.

9. Richard Edlund, *Cinefex* 100 (2005): 74.

10. Just as the real-world White House is covered with a special paint that diffuses light and prevents edges from gathering shadows. Caleb Kelly, in *Cracked Media: The Sound of Malfunction* (Cambridge: The MIT Press, 2009), shows how dirt, in forms like background noise, can signify realism in sound too. The genre of dirty noise in "music" involves cracking vinyl records, CD's, and turntables, to perform sound events. At the limit, though, such "redirection from signal to noise" (215) removes the sense of reference to the real.

11. Manovich (*Software*, 130) argues very little in software media is "born digital," without any reference to analogue media and/or physical objects. The interesting cases are like the "Wind" filter in photoshop, which creates tiny jagged lines on an image that makes it look windblown. At the micro level, the jagged lines do not look at all like windblown objects; at the macro level, it does.

12. Or the way people remix music videos with their own footage, or the way Duchamp's Green Box encouraged readers to make their own cut-ups with his texts.

13. Roland Barthes, *Elements of Semiology*, trans. Annette Lavers and Colin Smith (New York: Hill and Wang, 1983), 84.

14. In editing, we normally try to remove pops and hisses, by blocking high or low frequencies. If this is not possible in perception, we learn to live with them in different ways, as certain works of new music drive home. Everyone has thought of Smell-o-vision, but we tend to assume smell will contribute to the realism of the total art form (as Wagner imagined, and Bazin doubted). But we can equally wonder what special effects wizards will eventually do with odour?

15. Roland Barthes, *Empire of Signs*, trans. Richard Howard (New York: Hill and Wang, 1982), 14. A shimmering signifier is a signifier that shows almost nothing but the process of its composition, e.g. the Japanese dinner tray, containing food not cooked but only cut up, shifted around, and arranged in a clear, almost transparent medium such as soup.

16. Béla Balász, *Theory of the Film: Character and Growth of a New Art* (1945), trans. Edith Bone (New York: Dover Publications, 1970). Once perceivers know about image-effects, as Balász leads us to hope, then even when perceiving an image that somebody else has effected, they will experience the backgrounds and foregrounds that they know are in the graphic design. I cite Balász on a number of points, though other authors say similar things. One reason is that on many points, Balász said it first. More important, since Balász emphasized the importance of developing a systematic film theory for the first time, his enumeration of the topics in such a theory is useful for an attempt today to develop a systematic theory of visual effects.

17. Besides aesthetic objections, there might be objections based on principles of economic justice. For example, there might be an objection to visual effects editing on grounds that we generally have to buy programs from proprietary for-profit companies. But is that different from the way no scholar can do without Google and its ads? Or the way we have to buy eyeglasses, or house paint and clothing to colour the world?

18. One difference is that consciousness does not have an interface with the world. That is the point of phenomenology: to the things themselves. Another difference is that consciousness relies on pre-conscious syntheses and heuristics, whereas in editing, heuristics have to be made into conscious trouble-shooting devices. Final Cut typically anticipates the user's intention without direct instruction, whereas Avid requires the user to make every decision explicitly. Each program type has advantages and disadvantages. Embodied consciousness and editing programs do have one thing in common: they malfunction when they get old.

19. Kevin Mack, Visual Effects Supervisor, *Cinefex* 100 (2005): 64.

20. John Knoll, *Cinefex* 100 (2005): 24.

21. This program was developed for *The Polar Express*, 2004 (*Cinefex* 100 (2005): 125). To be sure, both representational painters like Meissoniet and impressionists like Degas studied the positions of horses' feet to get each framing right, but it is easier to know what is in each frame than to know programmatically how each position follows from the last.

22. The crew called it the Barrel-of-Monkeys scene.

23. Motion capture blurs the distinction between recording and drawing. A performer's movements are copied into an animation program, which is then animated; which is then sometimes followed by face replacement, pasting her recorded face over the animated body.

24. Tom Sito, *Moving Innovation: A History of Computer Animation* (Cambridge: The MIT Press, 2013), 214. I do not know if the Academy is thinking along these lines, but perhaps the distinction could be made not on the basis of whether the figures are recorded live, but on the basis of layering structure. Traditional drawn animation was able to use around six layers of coloring; the acetate images were 0.003 mm thick, so layering more than six sheets to photograph in one would lead to blurring backgrounds. There is no limit to the number of digital layers that can be blended without loss of content (Sito, *Moving Innovation*, 251).

25. And indeed whether an effect intensifies perception or de-saturates it.

26. Since perceptions are genetic, and its parameters interact in something like a syntagm, unintended consequences often interfere with what one thought one was about to perceive.

27. Furthermore, we cannot entirely separate perception and understanding. Editors still tend to use the doctrine that editing should "tell a story." The idea that there is a "language" of film, particularly a language of editing, had some initial appeal. Shot-reverse shot editing looked like a kind of subject-object grammar; zooming to a close-up looked like adjectival description. Though the idea of film grammar has been criticized in convincing detail, there remains some potential in Barthes' semiology of extra-linguistic systems.

28. Edgar Burcksen, "Capturing Cinematography," *CinemaEditor Qtr* V 63.2 (2013): 16-17.

29. Balász is concerned that when the camera can see things a "healthy man" cannot, like inside his own stomach (102-103), it is science but not art, since such a perspective cannot be that of a conscious subject. But by now we have seen enough stomach-innards scenes to know they can be subjective too.

30. In an analogous way, classical music composition assistance programs use scores on a staff, whereas IR-CAM's OpenMusic composition program for electro-acoustic music uses patches.

31. In other cases, two programs for the same kind of effect will produce different resulting images, for example generating fire using a particle program may look different from fire generated with a layer texturing program.

32. Alain Berthoz, *Emotion and Reason: The Cognitive Neuroscience of Decision Making*, trans. Giselle Weiss (Oxford: Oxford University Press, 2006), 137-71.

33. Balász, *Theory*, 41.

34. What is at stake is not sensing what is inside the organ (the way one hears the songs one sings inside one's closed mouth). Here, one controls one's organ from the screen outside one's body.

35. Layers in Photoshop were introduced in 1994. Photoshop Help: "Layers allow you to work on one element of an image without disturbing the others" (quoted by Manovich, *Software*, 142), and this has general implications for what an image "means" (142). By opening each frame for micro-control on the small scale (285), the effect of layering is to define all sorts of images by a single "metamedium" (145).

36. Balász, *Theory*, 30-31.

37. Transparency and "negative space," the place where the image is not, is not empty, but full space in its own right.

38. For example, the first audiences to see close-ups on screen thought they were severed heads, but we got used to seeing in scale.

39. This was discussed by Godard in 1959. Jean-Luc Godard, *Godard on Godard*, trans. and ed. Tom Milne (New York: Da Capo Press, 1972), 139.

40. What would human perception look like if we learned to squint by layers?

41. Balász emphasizes that in film viewing, the viewer has to add his own inferences. For example, the viewer sees lights flash against a man's face in a train station, and has to infer that the train has started to move and that it is the movement of its window lights that is visible on the man's face (36).

42. Audio for decades has worked with many tracks (or harmonic layers). But in some specifics of simultaneity-management, the eye is ahead of the ear, largely because while sounds have momentum, echoes, dissolves, and asynchronics, they do not have sides, angles, or shadows. Balász says this explains the problem faced by radio plays (Balász, *Theory*, 213-217). Like Deleuze (*Cinema 2: The Time-Image*, trans. Hugh Tomlinson (Minneapolis: University of Minnesota Press, 1994), 234-241), Balász (*Theory*, 180) highlights the staggering of audio and video "counterpoint." Layering does not cancel the difference between sight and sound. When video shows irrational images and the audio explains it rationally in words; or when in horror films the ear is given a sound and the camera eye then has to look for its source, layering staggers perception across media, across sense and source, reason and unreason. Here we are not talking about translating visual fluctuations to make different sounds (like oscilloscope imaging, the oldest form of electronic graphics), but the layering of meanings. Layering overrules the distinction between double system filmmaking, where video and audio are recorded by different devices and synchronized later, and single system recording with microphones inside a video camera. On the one hand, in compositing, it is as though every shot is recorded in multiple system; on the other hand, time codes can be matched automatically, so it is as though a single wired system networks across difference devices. It is as if all movies are now dubbed.

43. The latter two are discussed in Manovich, *Software*, 260-266. Work of all three can be found on YouTube.

44. The revolution in inexpensive video cameras in the 1970's, has been succeeded over the last fifteen years by a revolution in inexpensive digital editing, giving amateurs enormous access to image-control power.

45. Manovich takes the position that opening each frame re-defines images as "composition-based" *rather than* "time-based" (*Software*, 282-289), but the cases I raise blur this distinction.

46. Brown, *Supercinema*, 42-46.

47. Or frames left out to save disk space. Audio editors use the fact that we notice more changes in high frequencies than low: by sampling fewer pitch changes at the low frequencies, they save disk space with no loss of perception to humans.

48. Jody Duncan, "Zombie Wars," *Cinefex* 135, October 2013: 12-31.

49. There are many ways to control the curve, like the Bezier curves used to control spatial lines in Photoshop.

50. John Alton, *Painting with Light* (Berkeley: University of California Press, 1995).

51. Trish and Chris Meyer, *Creating Motion Graphics with After Effects*, fifth edition (Burlington, MA: Focal Press, 2010), 132.

52. Robert L. Carringer, *The Making of Citizen Kane* (Berkeley: The University of California Press, 1985), 82.

53. Jacques Lacan, *The Four Fundamental Concepts of Psycho-Analysis*, trans. Alan Sheridan (New York: W. W. Norton, 1981), 75.

54. Of course, a backside is not a Lacanian unconscious; computer graphics do not represent the desire of the other.

55. In controlled perception, the zero point of orientation is not tied to the balanced organism standing upright.

56. To compare art forms' treatments of backgrounds, compare the tree bark in *Shrek*, or the skin hematomas on *Wolfman*'s (2010) face, with Mantegna's way of painting the marble bases of statues.