

# Perceiving Causation via Videomicroscopy

Megan Delehanty<sup>†</sup>

---

Although scientific images have begun to receive significant attention from philosophers, one type of image has thus far been ignored: moving images. As techniques such as live cell imaging and videomicroscopy are becoming increasingly important in many areas of biology, however, this oversight needs to be corrected. Biologists often claim that there are relevant differences between video and static images. Most interesting is the idea that video images allow us to see causal relationships. By identifying the conditions that would be required for this to be true and showing that they are not satisfied at the micro level, I will show that videomicroscopy does not provide us with special access to causal information.

---

**1. Introduction.** Scientific images have begun to receive an increasing amount of attention from philosophers (e.g., Griesemer 1991; Ruse 1991; Hammer 1995; Kitcher and Varzi 2000; Perini 2005), but there is one category of images which has thus far been ignored: moving images. If there were no philosophically interesting difference between static and moving images (video), this oversight would not matter, but two things stand in the way of this conclusion. First, scientists frequently distinguish between video and static images, claiming that video provides us with more and better information, and in particular, allows us to directly perceive causal relationships. Second, there is a considerable literature in psychology, following on the work of Michotte ([1946] 1963), which suggests that humans *can* perceive causation. This work is based on having subjects view various sorts of animations and does not apply to static images. Thus, the idea that there may be relevant differences between types of data display format—specifically that video images alone allow us access to causal information—needs careful analysis.

<sup>†</sup>To contact the author, please write to: Department of Philosophy, University of Calgary, 2500 University Dr. NW, Calgary, Alberta T2N 1N4; e-mail: mdelehan@ucalgary.ca.

Philosophy of Science, 74 (December 2007) pp. 996–1006. 0031-8248/2007/7405-0037\$10.00  
Copyright 2007 by the Philosophy of Science Association. All rights reserved.

There are two ways in which we might interpret claims that video images allow us to *see causation*:

1. **LIT**: In viewing video images, we literally see causal relationships.
2. **ID**: In viewing video images, we are able to distinguish between different causal hypotheses that we could not distinguish on the basis of other data display formats.

While **ID**, if true, would certainly support the claim that video often provides an epistemic advantage over static images, it is **LIT** that is more philosophically interesting. In order to assess which interpretation(s) can be supported, it is necessary to make several distinctions. First, we need to separate features attributable to the data display format from those belonging to the experimental system as a whole. Techniques whose output is video often involve numerous differences from those whose output is static images, so to defend the claim that video confers an epistemic advantage, we need to eliminate other sources of epistemic difference. Second, we need to distinguish between video of objects at the micro level (that of cells and subcellular objects, events and processes), and video of objects at the macro level of everyday objects. In making claims about the advantages of video, biologists are normally referring to videomicroscopy or other techniques that involve visualizing events at the micro level. This will turn out to be important as I will argue that we can answer the question of whether video allows us to see micro-level causes without having to settle the thorny issue at the macro level. Finally, we will need to divide accounts of causation into two broad types: *singularist accounts* (**S**) that take causation as a relation intrinsic to a given interaction, and *regularity accounts* (**R**) that take causation to be an extrinsic relation, the identification of which requires background information in addition to the observed sequence of events. Both require that we acquire spatiotemporal information via observation, but on **S**, causal relations are literally observable, while on **R** they can only be inferred from what is observed.

In this paper, I will argue that neither **LIT** nor **ID** can be defended in the context of videomicroscopy. The basic outline of the argument is as follows: To defend **ID**, we need to show that video confers an advantage over static images on the remaining account **R**, and that this advantage comes from the data display format rather than from other features of the experimental system. Although video often does offer the sort of advantage claimed by **ID**, I will show that this advantage cannot be attributed to differences in the data display format (between video and static images). To defend **LIT**, **S** must be shown at least not to be clearly false. But **S** must be false at the micro level since (1) the requirements for perceiving Michotte-type causal interactions are rarely met and, even when they are, cannot be interpreted as actually causal without background

information, and (2) we do not know what causally efficacious actions at this level look like and must rely on background information obtained from other sources.

**2. Two General Accounts of Causation.** In order to determine whether or not we can get causal information from video but not from static images, we first need to have some idea of what we mean when we say that we *saw* X knock down Y or make some other sort of causal claim. Hume (1980) famously argued that causes are not knowable a priori and that the observation of regularities in the world serves as the basis for our impression of causality. All we really perceive, according to Hume, are constant conjunctions of objects or events. When we see one event regularly followed in space and time by another, the first one will naturally and forcefully bring to mind the expectation of the second. The causal impression is just this action of the mind. Although one can see the prior event that would be labeled ‘cause’ and the subsequent event that would be labeled the ‘effect’, it is not possible to see a causal connection between them. Other philosophers have contested this claim, maintaining that it is indeed possible to observe causation, even in single cases where no constant conjunction can be found. Thus, Ducasse (1994) argues that by observing the relata of a causal relation, we observe the cause. Anscombe (1994) claims instead that the concept of a causal relation is too abstract and has meaning only if we can first understand ideas like push, pull, break, bend, and so on.<sup>1</sup>

These causal concepts can be applied on the basis of observation—we know what it looks like for something to break or to push or pull another object. What we see, then, are instances of pushing, pulling, breaking, etc, not causation more generally.

Hume’s account is representative of the *regularity* type of account, **R**. Again, **R** holds that what we are doing when we ‘see’ causation is drawing an immediate, automatic inference based on our visual experience of this particular interaction together with background knowledge of some sort. Exactly what this background knowledge is supplying that allows us to identify certain relations as causal—information about regularities, counterfactual dependence, or something else—is far from agreed upon. But for the present purpose, disagreements of this sort can be passed over since the distinguishing feature of this position is that it holds that we cannot literally perceive causation. Anscombe’s and Ducasse’s accounts are representative of the *singularist* type of account, **S**. **S** claims that causal relations are immediately accessible to experience: we really are *seeing*

1. This type of account is also defended by Armstrong (1997), Cartwright (2000), Fales (1990), and Menzies (1998).

causation, just as we would see a color or a shape. Notice that for both **S** and **R**, there are two minimal conditions for ‘seeing’ an interaction as causal. The first condition is that we need to acquire spatiotemporal information about an interaction. This is provided by observation on either account. Where the two accounts differ is on how we identify an interaction as causal. On **S**, we directly see it as such on the basis of the observed event alone—nothing more is needed. On **R**, the causal attribution is made inferentially, on the basis of background information. While I cannot hope to resolve the question of which of the two alternatives more adequately describes what is happening when we make causal claims on the basis of seeing some interaction at the macro level, I will argue in Section 5 that **R** is the only possible alternative at the micro level of videomicroscopy. This is due to the fact that the careful application of background knowledge is essential to identifying actions such as phosphorylation and GTP (guanosine triphosphate) hydrolysis that possess, at the micro level, the causal efficacy that actions such as pushing and pulling have at the macro level. As a result of this, **LIT** must fail. However, as Section 5 will also show, static images can potentially provide us with the same spatiotemporal information as video images and so may allow us to *interpret* events to be causal on **R**. Thus, **ID** is not defeated by the falsity of **S** at the micro level but requires a further argument to establish that the epistemic advantage of video is not due to the data display format but rather to other differences in the experimental system usually associated with different data display formats. In Section 3, I will show what is required to isolate the effect of different data display formats from other differences in experimental systems. Sections 4 and 5 will then show that **S** must fail at the micro level and that static images can satisfy **R** as well as video images once the differences identified in Section Three have been eliminated.

**3. Epistemic Advantages of Video Images.** The advantages of live cell imaging (data from which is usually captured by videomicroscopy) have been widely celebrated by biologists though it is often unclear precisely what the advantage is supposed to be:

Being able to observe processes as they happen within the cell by light microscopy adds a vital *extra dimension* to our understanding of cell function. (Stephens and Allan 2003, 82; my emphasis)

With the advances in labeling and imaging technologies, we have already witnessed remarkable improvements in our ability to monitor and *interpret* processes in real life and in real time. (Hurtley and Helmuth 2003, 75; my emphasis)

The above quotations are from articles in a special issue of *Science* devoted to biological imaging. Live cell imaging is identified as giving us more information, but what is the significance of this extra information? The obvious response is that it gives us temporal information, but temporal information is not absent from all data presented as static images: series of static images produced at defined temporal intervals also convey this information. So what is it that we get exclusively from video? Though the review papers cited above do not make any direct reference to causal information, when we turn to reports of specific imaging studies, causal claims are widespread. Thus, for instance, we read

that kinetochores can attach to the forming spindle by *capturing* astral MTs [microtubules] was *directly demonstrated* by video microscopy . . . Subsequent video microscopy studies revealed that this kinetochore *switches* between two activity states: one that allows it to *move poleward in response to a force*, and another that allows it to be *pushed (or pulled) away*. (Rieder and Khodjakov 2003, 93; my emphasis)

These causal claims are, as above, often explicitly based on the interactions that are seen in the videos. Is it the case, then, that watching videos produced by imaging living cells allow us to identify causal relationships in a way that we cannot by viewing static images?

The problem with comparing series of static images to video is that they usually involve not only different forms of data display, but different experimental systems. There are at least three epistemically relevant types of difference. First, video normally permits a much greater temporal resolution. Second, sampling requires that each image in the time series represents different individual cells/molecules while live cell imaging can continuously monitor a single cell or molecule from start to finish of the imaged process. Third, for many objects and events there may exist no way of monitoring them or making them visible by any technique other than those normally used for live cell imaging. These are all clear advantages to video imaging, so in order to assess whether video images *also* have the advantage of giving us access to causal information, we need to compare images with the same data content. Fortunately, in the context of most modern imaging technologies, it is easy to see how this is possible. Numerical data is produced prior to the image (static or video) in all technologies where data received by the detection system is digitized before the production of the output image. Thus, there need not be any difference in content between any of these forms of data display (even if there usually is in practice). This is important since it means that we can isolate differences that are due to features of the data display format (static vs. video images) from those that must be attributed to differences

in the data collected by the imaging system. For the remainder of this paper, then, I will consider only comparisons drawn between images with a different format, but the same content. This should not be taken to deny the importance that other features of an experimental system have for the value of the data, but simply to recognize the fact that these virtues cannot be attributed to the form in which the data is displayed.

**4. Psychologists on Perceiving Causation.** The possibility that there is something intrinsic to an interaction in virtue of which we can *see* causation has been the subject of extensive investigation by psychologists. Not content to make empirical claims about human perception, some authors suggest that this work tells us something about what causation *is*. Thus, Scholl and Nakayama write: “while ‘causation’ is typically thought of as a high-level conceptual property, numerous experiments suggest that the visual system may itself traffic in causality” (2004, 455). This work also often explicitly makes reference to—and rejects—Hume’s claim that we cannot see causation. Twardy and Bingham, in a paper which uses Dowe’s conserved quantities (CQ) account of causation to help identify which perceptible properties of events are required to identify causal relationships, begin by suggesting that “if physical causal interactions are just exchanges of CQs, then in perceiving events, observers also perceive some important aspects of causation itself, contrary to Hume” (2002, 957). They proceed, in the course of the paper, to defend the CQ account and so derive the conclusion that human observers *do* perceive causation by perceiving exchanges of physical quantities. While I specifically want to avoid defending a specific account of causation here, this work might be taken to lend support to **LIT**, the claim that video allows us to literally see causation (however defined). What I will argue in this section, then, is that it ought not.

Since the work of the French psychologist, Albert Michotte, in the middle of the twentieth century, a large amount of research has been undertaken to investigate when the visual system will interpret a dynamic stimulus as causal (e.g., Michotte [1946] 1963; Scholl and Nakayama 2004; Twardy and Bingham 2002; White and Milne 2003). This work involves experiments such as the following (see Figure 1): Two shapes, **A** and **B**, are displayed and animated on a computer screen. Shape **A** begins to move towards **B** then stops when it is immediately adjacent to **B**. Just when **A** stops, **B** begins to move away from **A**. Observers are asked whether or not **A** was the cause of **B**’s motion. In the situation just described, the majority of observers will claim that **A** was the cause. However, if the setup is changed very slightly so that **B** starts moving a fraction of a second earlier or later or if **A** overlaps **B** before **B** starts moving, then the proportion of people who claim that **A** caused **B** to move drops

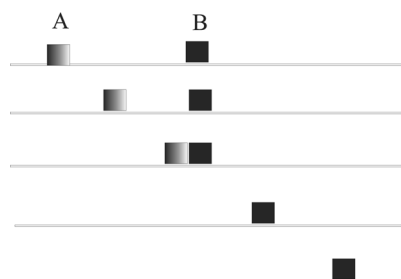


Figure 1. Illustration of Michotte's launching effect.

significantly. Michotte referred to the 'illusion' of causality in this interaction as the 'launching effect'. Other sorts of interactions have also been shown, by Michotte and later researchers, to produce the impression of causation.

The most important thing to notice about these experiments is that they indicate that there are tight constraints on the sorts of interactions, whether viewed only once or repeatedly, that humans perceive as causal. Some visual interactions almost always, seemingly unavoidably, produce the impression that they are causal, while others, no matter how regularly they occur, never do. This suggests that even if scientists can directly 'see' causation in videos of living cells, they will do so only for a limited subset of all actual causal interactions. The animations used in these experiments are extremely simple, even when contextual factors are added to see how they influence the perception of causation. Cases of observed interactions in, for instance, a confocal microscope are virtually always far more complex than the simple interaction just described. There are many more objects and the types of interactions will not often fit the limited set of spatiotemporal conditions under which we unavoidably 'see' causation. So the visual interactions observed via videomicroscopy will almost never produce a causal impression in the sense described by Michotte and others. Moreover, even if such a straightforward type of interaction were to be observed, there is no reason to think that the causal impression corresponds to any actual causal relation. After all, no actual causation is involved in the animations used by psychologists. We may 'see' causation where it fails to exist and fail to 'see' it where it is present. Many cases where we may have good reason to say that a protein-protein interaction involves some sort of causal activity such as phosphorylation, for instance, will not involve the sorts of visual interactions that people identify as causal (e.g., launches) and, even if they do, they will not fall within the correct spatiotemporal boundaries.

The conclusion that we cannot ‘see’ causation in this simple sense in video images is not enough to establish that we cannot get causal information from video (or from other forms of data). A Michotte-type ‘seeing’ would provide one possible defense of **LIT**, but it is not the only—or best—alternative.<sup>2</sup> In order to defeat **LIT**, it is still necessary to look at the sort of information that seems to be involved when observers ‘see’ certain interactions as causal. I suggested earlier that the information that both **R** and **S** require us to get from the data itself is spatiotemporal relationships between objects. But we also need additional information to *interpret* (rather than simply ‘see’ in the Michotte sense) an interaction as causal whether or not it is also ‘seen’ as causal. This information is background knowledge of the type of mechanism that is plausible in a given context.

**5. What Are We Seeing When We ‘See Causation’ at the Micro Level?** It is unclear exactly what is meant by the term ‘causal’ in the animation experiments described above. Michotte simply asked observers how ‘causal’ an interaction seemed to be. What must be involved, however, are certain types of spatiotemporal relationships. A launch event is perceived, for instance, if a moving object gets close to another then stops, and the second object, after a suitably small time interval, begins to move in a certain direction with a suitable velocity. What counts as ‘suitable’ in these instances is, presumably, determined by some part of the human visual system or other cognitive apparatus.

It may be that a similar story about our ability to recognize pushing, pulling, tearing, etc., could be told. Some such account, at least, is required to establish the plausibility of **S**. This may or may not be possible at the macro level. (Can we really distinguish if John fell over because I pushed him or because he faked falling over in response to my hand approaching him but never quite making contact?) But whatever may be the case at the macro level, the difficulty, when it comes to identifying causal activities at the micro level, is that we cannot get sufficiently fine-grained spatiotemporal information to recognize causal concepts like phosphorylation. We are unable to give precise descriptions of the spatiotemporal characteristics of most of the sorts of events or processes that we want to say are causally responsible for some change. ‘Pushing’ and ‘pulling’ may describe what a micro level interaction looks like to us, but these terms do not represent actions to which we normally attribute causal efficacy at this level. Rather, what we are concerned with are actions such as

2. Whether Michotte events involve seeing causation or merely interpreting events as causal would need to be addressed were this data to be taken as supporting the idea that we can actually perceive causation.



phosphorylation or GTP hydrolysis. But if some protein, A, phosphorylates another protein, B, which, once phosphorylated, undergoes a conformational change and dissociates from some third protein, C, all we will likely see is that A made contact with B and then B moved away from C. What is involved in claiming that A phosphorylated B, therefore, is not analogous to seeing the pushing or pulling that Anscombe claims is at the root of our observation of causes. Additionally, we don't know what it looks like for A to phosphorylate B in the same way that we know what it looks like for A to 'launch' B or for one person to hit another. And even if we did have this knowledge, the resolution of most of our imaging technologies is insufficient to allow us to discriminate between different causes on the basis of their appearance alone. An interaction between a GTP-binding protein and a GTPase activating protein (GAP) that causes the hydrolysis of GTP may well look just the same as some kinase A phosphorylating protein B in a confocal video. What we can observe is the changing spatiotemporal relationship between the two proteins and other parts of the cell, not the supposedly causal relation (GTP hydrolysis or phosphorylation) itself. To determine which of these processes is actually occurring requires additional information about which proteins have been labeled and what sorts of activities they may engage in. This, however, is not information that is present in the imaging data, whatever format is used to display it. Thus, the ineliminable role of background information in interpreting events as causal means that **S** must be false at the micro level.

Background information is required to identify causal relations at the micro level in any data display format. In most cases of biological imaging, we will not get a Michotte style causal impression from a video (and it is in principle impossible for us to do so from other data formats). Whether or not we 'see' a causal relation in the data, in order to identify an actual causal relationship we must have background information about the types of interactions that particular objects, in particular sorts of spatiotemporal relationships to one another, can or cannot participate in. Background information insofar as it supplies the possibility of a plausible mechanism for a causal interaction, is required to produce more than a descriptive statement about the spatiotemporal positions of the objects under investigation. A launch or other interaction may be *seen* as potentially causal, but can be *interpreted* as actually causal only provided that there is a mechanism that identifies the smaller scale causal concept (phosphorylation, etc.) and can explain this interaction as causal. The same sort of information is required if the interaction is *not* seen as causal. Given the severe constraints on the sorts of interactions that we 'see' as causal, very few biological interactions will be seen as causal, but this has no impact on whether or not they can be *interpreted* as causal.

One role for the background information is to identify the (possible) small scale causes actually involved in a larger scale (noncausal) interaction such as A approaching or moving away from B. Another is to rule out possible alternative explanations of some event, resulting in the possibility that an actual (low temporal resolution) time series of static images could be epistemically equivalent to a video image when the set of possible causal events or interactions occur at a time scale greater than the interval between images in the series. Just as there are constraints on the spatio-temporal conditions under which we will 'see' a launch or other causal event, background information places upper and lower bounds on the larger scale spatial and temporal relations (those that we actually observe) that can be connected to the smaller scale causal interactions such as phosphorylation that need to be either ruled out or permitted. It is not necessary, for instance, that for one object to be claimed to have caused another's motion, that there be no temporal gap. If phosphorylation or some sort of conformational change is supposed to be initiated by the arrival of A close to B and responsible for causing B to start to move away, it is entirely reasonable to expect that there will be a gap between the arrival of A and the departure of B.

Although Section 3 showed that many of the differences between video and static images are due to differences in the associated experimental systems, **ID** could still be true if spatiotemporal information of the sort required by **R** is only present in video images. It should be obvious, however, that information about the relative spatiotemporal positions of various objects is present not only in video images, but also in series of static images.<sup>3</sup> If we had a time series of static images at intervals equal to the inverse of the frame collection rate of the video camera, the two display formats would contain exactly the same spatiotemporal information (though in practice, the temporal resolution for the time series will usually be much lower). Since the background information required, on **R**, to supplement the spatiotemporal information is not derived from the image (static or video), there is no difference between data display format with respect to the identification of causal interactions.

**6. Conclusion.** The data that is acquired by many biological imaging technologies can be presented in different formats such as static images, as videos, as graphs, as diagrams, or even as very large sets of numbers.

3. The same information is also present in numerical data. This information can be extracted; though with more difficulty, since the ability to recognize an object (in particular, its boundaries) is aided by visual presentation. However, one could use computational methods to identify objects (as they move and change over time) in the numerical data.

Video images, however, are often claimed to have the advantage of providing us with causal information. While live cell imaging does often get us more information—and sometimes causal information—than we can get from series of static images, this is due to the kinds of intervention that different imaging technologies allow rather than to an epistemic difference between video and static images. To claim otherwise would require either that we be able to immediately recognize micro level causes or that static images are not be able to provide temporal information. Neither of these is the case, thus videomicroscopy does not allow us special access to causal information.

## REFERENCES

- Anscombe, G. E. M. (1994), “Causality and Determination” in E. Sosa and M. Tooley (eds.), *Causation*. New York: Oxford University Press, 88–104.
- Armstrong, D. M. (1997), *A World of States of Affairs*. Cambridge: Cambridge University Press.
- Cartwright, N. (2000), “An Empiricist Defence of Singular Causes”, in R. Teichman (ed.), *Logic, Cause, and Action*. Cambridge: Cambridge University Press, 47–58.
- Ducasse, C. J. (1994), “On the Nature and Observability of the Causal Relation”, in E. Sosa and M. Tooley (eds.), *Causation*. New York: Oxford University Press, 125–136.
- Fales, E. (1990), *Causation and Universals*. New York: Routledge.
- Griesemer, J. R. (1991), “Material Models in Biology”, in A. Fine, M. Forbes, and L. Wessels (eds.), *PSA 1990*, vol. 2. East Lansing, MI: Philosophy of Science Association, 79–93.
- Hammer, E. (1995), *Logic and Visual Information*. Stanford, CA: CSLI Publications.
- Hume, D. (1980), *A Treatise of Human Nature*. 2nd rev. ed. Oxford: Clarendon.
- Hurltley, S. M., and L. Helmuth (2003), “The Future Looks Bright”, *Science* 300 (5616): 75.
- Kitcher, P., and A. Varzi (2000), “Some Pictures Are Worth 2<sup>80</sup> Sentences”, *Philosophy* 75: 277–381.
- Menzies, P. (1998), “How Justified Are the Humean Doubts about Intrinsic Causal Links?”, *Communication and Cognition* 31: 339–364.
- Michotte, A. ([1946] 1963), *The Perception of Causality*. Trans. T. Miles and E. Miles. New York: Basic Books. Originally published as A. Michotte (1946), *La perception de la causalité*. Louvain: Publications Universitaires.
- Perini, L. (2005), “The Truth in Pictures”, *Philosophy of Science* 72: 262–285.
- Rieder, C. L., and A. Khodjakov (2003), “Mitosis through the Microscope: Advances in Seeing inside Live Dividing Cells”, *Science* 300: 91–96.
- Ruse, M. (1991), “Are Pictures Really Necessary? The Case of Sewall Wright’s ‘Adaptive Landscapes’”, in A. Fine, M. Forbes, and L. Wessels (eds.), *PSA 1990*, vol. 2. East Lansing, MI: Philosophy of Science Association, 63–77.
- Scholl, Brian J., and Ken Nakayama (2004), “Illusory Causal Crescents: Misperceived Spatial Relations Due to Perceived Causality”, *Perception* 33: 455–469.
- Stephens, D. J., and V. J. Allan (2003), “Light Microscopy Techniques for Live Cell Imaging”, *Science* 300: 82–86.
- Twardy, C. R., and G. P. Bingham (2002), “Causation, Causal Perception, and Conservation Laws”, *Perception and Psychophysics* 64: 956–968.
- White, P. A., and A. Milne (2003), “Visual Impressions of Penetration in the Perception of Objects in Motion”, *Visual Cognition* 10 (5): 605–619.