The perception of causality: Feature binding in interacting objects

John K. Kruschke and **Michael M. Fragassi**
Dept. of Psychology
Indiana University
Bloomington, IN 47405
kruschke@indiana.edu
http://www.indiana.edu/~kruschke/home.html

Abstract

When one billiard ball strikes and launches another, most observers report seeing the first ball *cause* the second ball to move. Michotte (1963) argued that the essence of phenomenal causality is "ampliation" of movement, in which the motion of the first object is perceptually transferred to the second object. Michotte provided only phenomenological evidence, however. We extend the reviewing paradigm of Kahneman, Treisman, and Gibbs (1992) to Michotte-style launching events and report response-time data consistent with Michotte's notion of ampliation. We discuss how contemporary theories of feature binding can extend to the domain of interacting objects and address our results. We also suggest that our treatment of ampliation helps clarify controversies regarding whether perceived causality is direct or interpreted and whether it is innate or learned.

Causality Perceived

Imagine a billiard table. The cue ball rolls across the felt and strikes the 8-ball, launching the 8-ball. Most observers will report seeing the cue ball *cause* the 8-ball to move. The perception of causality has been placed at the foundation of cognition: Kant (1788) argued that causality was an innate and fundamental category of cognition; Piaget (1971) made it an integral part of his theory of development; and, more recently Leyton (1992) has argued that the extraction of causality is at the core of perception and cognition.

The British empiricist philosopher David Hume (1739) argued that impressions of causality are mere fabrications of a sophisticated mind: In the case of the billiard balls, the observer sees the cue ball move, sees the contact of the balls, and sees the subsequent motion of the 8-ball, but does not see causality itself. The impression of causality, Hume argued, is a learned nexus from the first to second ball, based on repeated observations of the conjunction of the two motions, their spatio-temporal contiguity, and the temporal priority of the one motion relative to the other.

Two hundred years later, the Belgian psychologist Albert Michotte impugned Hume, arguing instead that the impression of causality is a spontaneous perceptual gestalt, which is neither learned nor an interpretation via abstract knowledge of physical events (Michotte, 1941, 1963). Michotte claimed that the essence of perceived causality is "ampliation of motion." The neologism, "ampliation," refers to two aspects of the perceived motion. First, the motion of the approaching object is *transferred* to the launched object. Second, for a brief time just after impact (approximately 200ms), the motion is phenomenologically duplicitous: It *belongs* to the first object while the second object *has* it. Thereafter, the motion of the second object becomes autonomous.

Unfortunately, Michotte made only phenomenological observations of perceived causality and gathered no performance data, leaving open the possibility that ampliation of motion is merely an idiosyncratic epiphenomenon (Boyle, 1972; Joynson, 1971). Nevertheless, Michotte's methods and findings are frequently described in textbooks on perception, development, artificial intelligence, etc. (e.g., Boden, 1977; Bower, 1982; Bruce & Green, 1990; Rock, 1975), and a compilation of newly translated articles by Michotte has recently been published (Thinès, Costall, & Butterworth, 1991). In this paper we report a response-time experiment that yielded results consistent with Michotte's theory of ampliation.

We propose that the theory of ampliation can be construed to impact directly on theories of feature integration; i.e., theories of how different visual features of an object, such as shape, color, and movement, are bound into a common identity but distinguished from visual features of other objects in the same field of view (e.g., Treisman, 1986). The key idea is that if the motion is perceptually transferred from one object to the next, then the feature of movement must be unbound from the launching object and bound with the launched object. We discuss how contemporary theories of feature binding can account for Michotte's phenomenology of perceived causality.

Providing a performance measure of ampliation and giving it a theoretical interpretation in terms of feature binding also supplies a new perspective on the relation between ampliation and perceived causality. Rather than debate whether a single process of perceiving causality is either innate or learned, we suggest that the sub-process of ampliation might be perceived directly and developed early in infancy, but the complete perception of causality might be interpreted and learned.

An Empirical Approach to Measuring Ampliation

The Reviewing Paradigm

The performance measure of ampliation that we will describe is an extension of the *reviewing paradigm* invented by Kahneman et al. (1992). In this paradigm, the observer is first shown two objects on a computer screen, such as the triangle

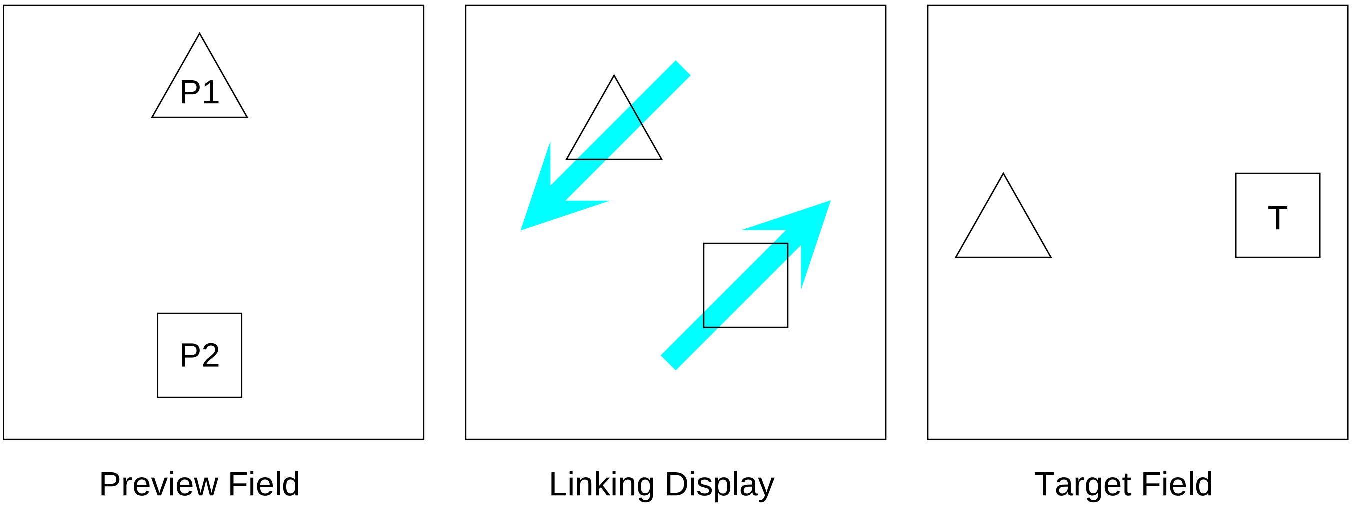

Figure 1: Example of the reviewing paradigm used by Kahneman et al. (1992).

and square in the left panel of Figure 1. Two letters, labeled "P1" and "P2" in Figure 1, are briefly presented ("previewed") inside each object. After the letters disappear, the empty objects visibly move to new locations, as shown in the middle panel of Figure 1. The motion of the objects links their initial positions to their final positions, and so this motion is called the "linking display." A target letter, labeled "T" in the right panel of Figure 1, then appears in one of the objects. The observer's task is to identify the letter as quickly as possible, by saying the letter's name into a microphone.

The key result is that observers can identify target letters that matched the preview letter from the *same* object faster than they can identify target letters that matched the preview letter from the *other* object. In Figure 1, for example, observers are faster to identify the target, T, when it matches P2 (which was in the same object as the target) than when it matches P1 (which was in the other object). Kahneman et al. (1992) called this effect the *object-specific preview advantage*.

The Reviewing Paradigm Applied to Launching

Kruschke (1987) applied the reviewing paradigm to Michotte-style launching events. Suppose the linking display in the reviewing paradigm did not keep the objects separated, as in Figure 1, but instead showed one object striking and launching the other, as one billiard ball can strike and launch another. Consider what would happen if the target letter appeared in the launched object. Would there still be a strong object-specific preview advantage, or would the preview information from the launching object be transferred, or ampliated, to the launched object? Kruschke (1987) reported that the object-specific preview advantage was significantly reduced in launching, relative to a control event in which the objects did not interact.

We replicated and extended that study in new experiments. In an event we call *launching*, shown in the left panel of Figure 2, participants saw one circle move toward and contact another, at which time the first circle stopped and the second circle moved away at the same velocity previously had by the first circle. Analogous to the reviewing paradigm of Kahneman et al. (1992), symbols such as "@" or "&" appeared briefly in the initial moments of the event, indicated by P1 and P2 in Figure 2. Then one empty circle launched the other, and a target symbol, indicated by T in Figure 2, appeared in the launched object, at which time the launched object was stationary. Unlike the identification task used in the experiments of Kahneman et al. (1992), the task for our participants was to indicate as quickly as possible whether the target symbol was the same as either of the two preview symbols. Responses were made by pressing a button when the target matched either of the preview symbols, and by pressing a different button when the target did not match either of the preview symbols.

A second event type, called *delayed motion*, began and ended the same way as launching, but had the two circles remain in contact with each other for approximately 890ms. According to Michotte (1963, Experiment 29, p. 91), observers perceive delayed motion as two independent movements without ampliation. The first circle is seen to stop completely, and the second circle then appears to move away with its own motion. Because the two motions are perceived as independent, we would expect to find a robust object-specific preview advantage in delayed motion. In launching, however, we predicted that the object-specific preview advantage would be diminished.

Participants also saw two other events in which the impacted object did not move. The third panel of Figure 2 shows the event we call *target at contact*, and the fourth panel shows *delayed target at contact*. In these events, the target appeared at the moment corresponding to when the first circle contacted the second circle in the launching event. The motivation for

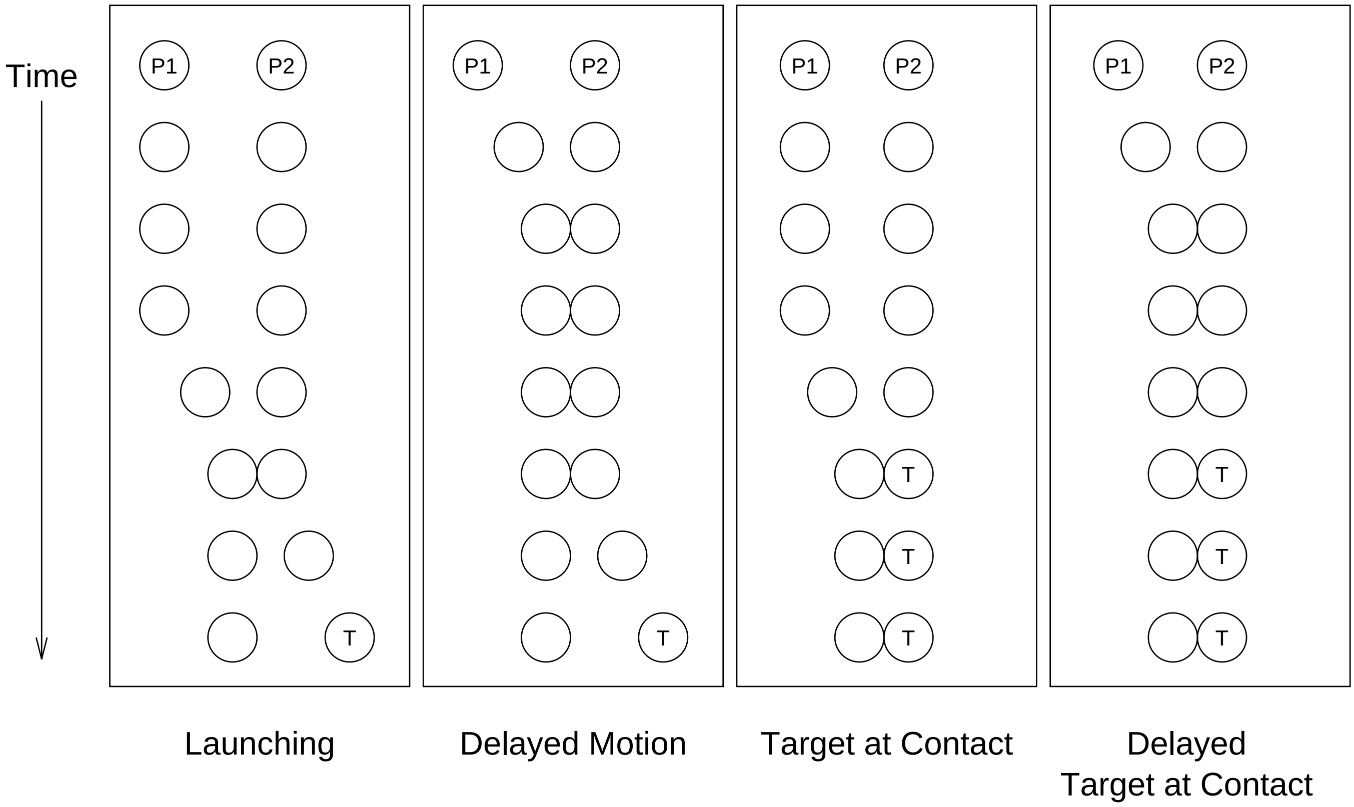

Figure 2: Schematic diagram of the four events in our experiment (not drawn to scale).

these events was to encourage observers to attend to the point of impact (which is where the target appears in these "contact" events). Previous experiments suggested that if observers see only the launching and delay events, they might immediately move their attention after the preview field to the anticipated location of the target, without devoting much attention to the motions of the objects. Attending to the point of impact was deemed important by Michotte for eliciting the best phenomenal ampliation; attending to the point of impact is also important to clearly distinguish launching from delayed motion. We anticipated a strong object-specific preview advantage for delayed target at contact, but a diminished advantage in target at contact.

Method

Participants. Forty-two undergraduates at Indiana University volunteered in partial fulfillment of a psychology course requirement.

Stimuli. Stimuli were presented on a PC-type 13" color monitor in VGA resolution. On every trial, there was a yellow rectangular frame, 256mm wide by 60mm high, which enclosed the relevant region of the the display. The circles had a diameter of approximately 27mm, and were separated center-to-center by approximately 38mm. Every trial began with the circles and their preview symbols in the same position, centered laterally on the screen with a small yellow fixation dot centered between them. One circle and symbol were red, the other green. The fixation dot and preview symbols appeared for 500ms. The preview symbols then disappeared, and the empty circles remained stationary for 200ms. For launching (see Figure 2), the circles remained stationary for an additional 890ms and then underwent the launching motion for 230ms. When moving, the speed of the circles was approximately 33cm/s. Then the target letter appeared in the launched object. When the target letter matched a preview letter, the target letter had the same color as the preview letter it matched. The other events had the same total duration, with only the linking display differing between them. For example, in delayed motion, the linking display consisted of approximately 115ms for the motion of the first object, followed by 890ms of contact, following by 115ms for the motion of the second object.

Design and procedure. The design consisted of five crossed factors: (1) type of match between target and preview letters (match same object, match other object, match neither object); (2) direction of motion (left, right); (3) color of the left preview symbol and circle (red, green); (4) time of initial movement (late as in launching and target at contact, early as in the delay events); and (5) time of target appearance (late as in launching, early as in contact). The trials in which

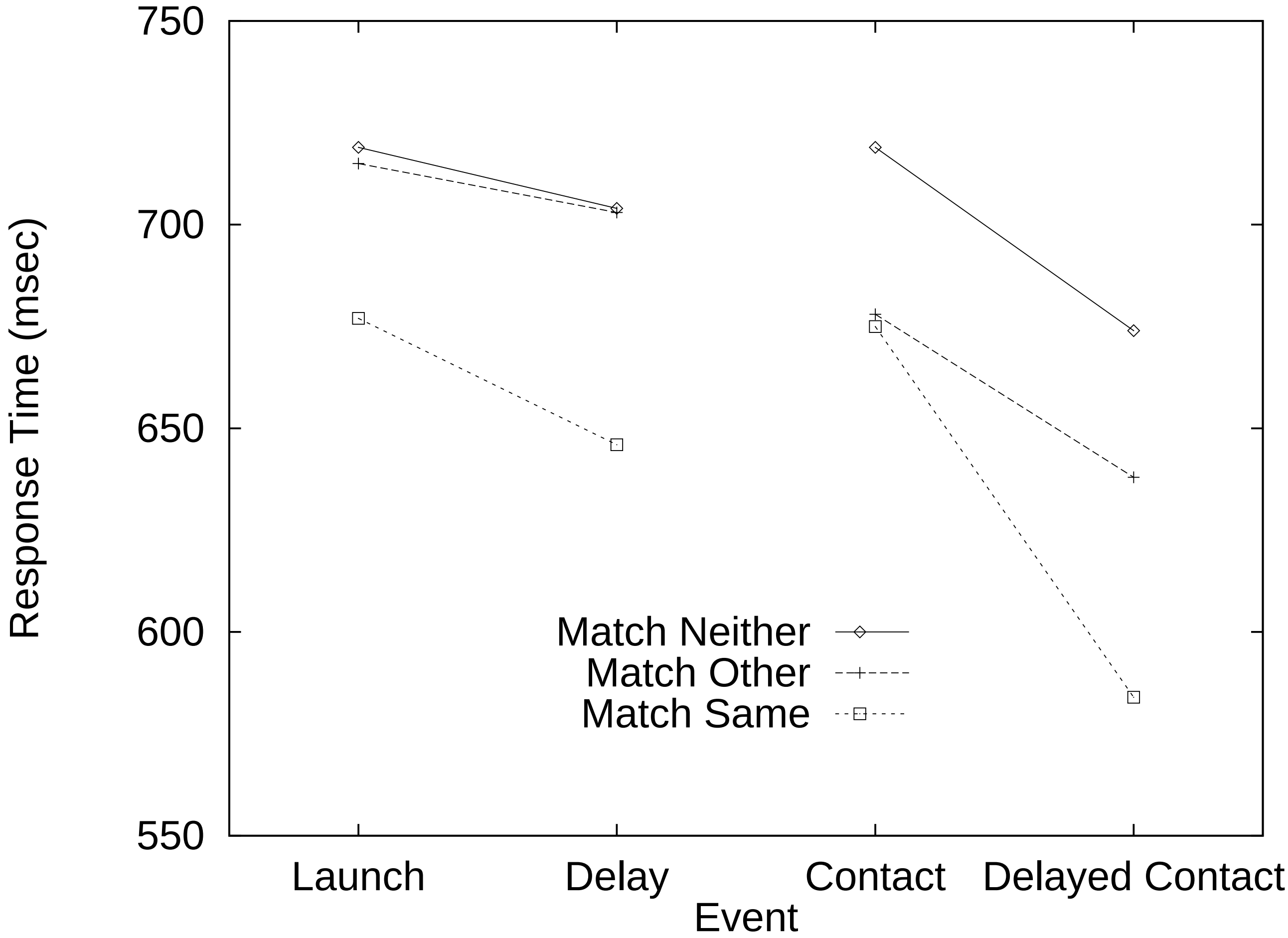

Figure 3: Mean response times for correct responses.

the target did not match a preview letter were doubled in order to equalize the total number of match and no-match trials. Each block consisted of 66 trials: 2 warm-up trials chosen at random from the design, followed by 64 trials that visited each cell of the design once.

After on-screen instructions with several examples of the events, participants had an initial practice block of 35 trials, followed by 8 blocks of 66 trials each, with brief rests between blocks. The experiment lasted about an hour.

Results

Figure 3 shows the mean response times as a function of event and match condition. These factors did not interact with direction of motion or color of symbol. There was indeed a strong object-specific preview advantage for delayed motion and for delayed target at contact, with response times significantly longer for matching the target to the other object than for matching the target to the same object (for delayed motion, RTs for match-other and match-same were 703ms vs. 646ms, respectively, $F(1,41) = 53.71$, $p = .0001$). The magnitude of the object-specific advantage was reduced for launching, although the reduction was statistically of marginal significance (interaction contrast $F(1,41) = 3.78$, $p = .059$). Comparable reductions have been observed in several other experiments conducted in our laboratory, and by Kruschke (1987), so we consider this trend toward reduction to be reliable. The reduction in object-specific preview advantage was quite strong for target at contact, however (interaction contrast $F(1,41) = 13.50$, $p = .0007$). Thus, at the time of impact, there was a complete loss of object-specific preview advantage, but moments later, in the delay events, the object-specific advantage was regained.

These results are intriguing for two reasons. First, regardless of their theoretical interpretation, the results show that the object-specific preview advantage is influenced by the type of interaction between the objects. In particular, mere *contact* of the objects, as in delayed motion, does *not* obliterate the object-specific preview advantage, but contact *with launching* does diminish the advantage. Moreover, the reduction in object-specific preview advantage is temporally localized just after the time of impact, and the object-specific advantage is regained moments later. Second, one interpretation of the results is that information from the launching object is ampliated to the launched object, so that just after impact the previewed symbol from the launching object is as accessible to retrieval as the previewed symbol from the launched object. These results are the first performance measure directly addressing Michotte's (1963) notion of ampliation in perceived causality.

Ampliation as Feature Re-Binding

To explain their results, Kahneman et al. (1992) suggested a "reviewing model," in which the motion of the linking display (middle panel of Figure 1) biases which object will be reviewed first in memory. The model assumes that the motion links the target object with the corresponding preview object, but the model does not incorporate any aspects of the dynamics of the motion. In particular, the model does not suggest why different events would have different biasing effects with different durations, as found in our data. Treisman and Kahneman (Treisman, 1986; Kahneman et al., 1992) have also suggested that visual features are bound together into "object files" that retain each object's immediate history. An object file is tantamount to a tag, or label, on each visual feature, identifying the object to which the feature belongs.

We propose that our results are mediated by binding and unbinding of the visual feature of motion with other features of the objects. One potential (and controversial) mechanism for feature integration is synchronization of neural pulses from different feature detectors. The pulse train acts as an object label for the feature, and synchronization gives the labels from different features a common signature. There is empirical evidence that neurons use this mechanism (for a recent review see Singer & Gray, 1995), and several researchers have applied this binding mechanism in models of visual perception, attention, and memory (e.g. Damasio, 1990; Grossberg & Sommers, 1991; Hummel & Biederman, 1992; Lumer, 1992; Mozer, Zemel, & Behrmann, 1992; Pabst, Reitboeck, & Eckhorn, 1989; Sporns, Tononi, & Edelman, 1991).

Applied to the scenario of launching, these ideas suggest that as the launching object approaches the to-be-launched object, the pulse trains of the motion detectors for the motion of the launching object are synchronized with the pulse trains of the other features of the launching object, all of which are desynchronized from the pulse trains of the features of the to-be-launched object. The question then becomes, What happens to the synchronization of the pulse trains when the objects come into contact?

One answer to this question is suggested by the hypothesis that motion primes other motion detectors along the forward trajectory. Long-range directional priming of motion detectors has been discussed extensively in models of visual motion perception developed by Marshall (1990; Martin & Marshall, 1993), and empirical evidence of long-range directional connections between motion-detecting neurons comes from work by Gabbott, Martin, and Whitteridge (1987) and others cited by Marshall (1990).

Our results might then be explained as follows: In the launching event, the two objects initially have desynchronized pulse trains. At impact, the directional priming of motion detectors causes the new movement of the launched object to have the same synchronization as the launching motion, so that just after impact the motion of the launched object is synchronized with the launching object, but localized with the launched object. This accounts for the duplicity of motion in Michotte's phenomenology: The motion of the launched object is localized with the launched object, but still belongs to (is synchronized with) the launching object. This also accounts for the difficulty and equalization of accessing the history of both objects: The target that appears in the launched object at the moment of impact is not clearly synchronized with either pulse train, and so both preview symbols are retrieved with equal difficulty. After a brief time, the motion of the launched object becomes synchronized with the other features of the launched object, and hence the object-specific preview advantage is resurrected.

This explanation is not committed to pulse train synchronization as the only possible binding mechanism. The explanation merely requires that the process of binding takes some small but non-zero time, and that the object tag of the launching object primes the launched object at the time of impact. Pulse train synchronization is just one possible mechanism for implementing these principles.

Ampliation versus Causality

Michotte (1941, 1963) argued that the perception of causality is not an interpretation based on acquired knowledge of mechanical events, but instead is perceived directly, with ampliation as its essence. Our results do not necessarily support this perspective, in full. It is possible that ampliation, *qua* feature unbinding and rebinding, is a direct perceptual mechanism, but the perception of causality is an additional interpretive process. For example, Weir (1978) described a model that classifies Michotte-style collision events without ever mentioning ampliation, presumably because of the concept's intangibility in Michotte's theories. Reification of ampliation, as suggested by our experiments, calls out for theories to address it, and provides one avenue for distilling which aspects of perceived causality are direct and which are interpreted.

Michotte also argued that the perception of causality is innate. Leslie (1982; Leslie & Keeble, 1987) provided evidence that six-month old infants can distinguish causal from non-causal events, or at least are sensitive to reversals of agency in causal events. Whether the infants perceive such events as causal or not remains an open question. It might be that ampliation develops rapidly, like stereopsis, in response to the visual world, whereas interpretations of causality are learned later in childhood. Perhaps ampliation is used as a perceptual cue for subsequent causal interpretation, and so the two are correlated. Separating ampliation from causality also allows for the possibility of universal sensitivity to ampliation, but individual differences in the perception of causality (Beasley, 1968; Schlottman & Anderson, 1993).

Acknowledgments

This research was supported in part by NIMH FIRST Award 1-R29-MH51572-01 to Kruschke, and by an Indiana University Cognitive Science Program Summer Research Fellowship to Fragassi. Thanks to Michael Erickson, Mark Johansen and Nathaniel Blair for comments on a previous draft, and to Colin Bogan, Amanda Reed, and Eddy Riou for administering experiments. Thanks also to Daniel Kahneman and Anne Treisman for encouragement and access to their lab at U.C. Berkeley in 1986-87.

References

Beasley, N. A. (1968). The extent of individual differences in the perception of causality. *Canadian Journal of Psychology*, *22*, 399–407.

Boden, M. A. (1977). *Artificial intelligence and natural man*. Basic Books, New York.

Bower, T. G. R. (1982). *Development in infancy*. W. H. Freeman, San Francisco.

Boyle, D. G. (1972). Michotte's ideas. *Bulletin of the British Psychological Society*, *25*, 89–91.

Bruce, V., & Green, P. R. (1990). *Visual perception: physiology, psychology & ecology*. Erlbaum, East Sussex, UK.

Damasio, A. R. (1990). Synchronous activation in multiple cortical regions: a mechanism for recall. *Seminars in the Neurosciences*, *2*, 287–296.

Gabbott, P. L. A., Martin, K. A. C., & Whitteridge, D. (1987). Connections between pyramidal neurons in layer 5 of cat visual cortex (area 17). *Journal of Comparative Neurology*, *259*, 364–381.

Grossberg, S., & Sommers, D. (1991). Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. *Neural Networks*, *4*, 453–466.

Hume, D. (1739). *A treatise of human nature*. Oxford, England: Oxford University Press. Republished 1980.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, *99*, 480–517.

Joynson, R. B. (1971). Michotte's experimental methods. *British Journal of Psychology*, *62*, 293–302.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*, 175–219.

Kant, I. (1788). *Critique of pure reason*. St. Martin's Press, New York. Translated by N. K. Smith, 1965.

Kruschke, J. K. (1987). The perception of causality: A performance measure of ampliation.. Paper presented at the Ninth Annual Berkeley-Stanford Conference in Cognitive Psychology. Available from the author.

Leslie, A. (1982). The perception of causality in infants. *Perception*, *11*, 173–186.

Leslie, A., & Keeble, S. (1987). Do six-month-old infants perceive causality?. *Cognition*, *25*, 265–288.

Leyton, M. (1992). *Symmetry, Causality, Mind*. MIT Press, Cambridge, MA.

Lumer, E. D. (1992). Selective attention to perceptual groups: The phase tracking mechanism. *International Journal of Neural Systems*, *3*, 1–17.

Marshall, J. A. (1990). Self-organizing neural networks for perception of visual motion. *Neural Networks*, *3*, 45–74.

Martin, K. E., & Marshall, J. A. (1993). Unsmearing visual motion: Development of long-range horizontal intrinsic connections. In Hanson, S. J., Cowan, J. D., & Giles, C. L. (Eds.), *Advances in Neural Information Processing Systems 5*. Morgan Kaufman.

Michotte, A. (1941). La causalité physique est-elle une donnée phénoménale?. *Tijdschrift voor Philosophie*, *3*, 290–328. Unpublished translation by Krista Hensley, April 30, 1993.

Michotte, A. (1963). *The perception of causality*. New York: Basic Books. Translated from the French by T. R. and E. Miles.

Mozer, M. C., Zemel, R. S., & Behrmann, M. (1992). Learning to segment images using dynamic feature binding. *Neural Computation*, *4*, 650–665.

Pabst, M., Reitboeck, H. J., & Eckhorn, R. (1989). A model of preattentive region definition based on texture analysis. In Cotterill, R. M. J. (Ed.), *Models of Brain Function*, pp. 137–150. Cambridge University Press, Cambridge, England.

Piaget, J. (1971). *Les Explications Causales*. Presses Universitaires de France, Paris. Translation by D. & M. Miles, *Understanding Causality*. New York: Norton. 1974.

Rock, I. (1975). *An introduction to perception*. Macmillan, New York.

Schlottman, A., & Anderson, N. H. (1993). An information integration approach to phenomenal causality. *Memory & Cognition*, *21*, 785–801.

Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, *18*, 555–586.

Sporns, O., Tononi, G., & Edelman, G. M. (1991). Modeling perceptual grouping and figure-ground segregation by means of active reentrant connections. *Proceedings of the National Academy of Sciences USA*, *88*, 129–133.

Thinès, G., Costall, A., & Butterworth, G. (Eds.). (1991). *Michotte's experimental phenomenology of perception*. Erlbaum, Hillsdale, NJ.

Treisman, A. M. (1986). Features and objects in visual processing. *Scientific American*, *255*(5), 114B–125.

Weir, S. (1978). The perception of motion: Michotte revisited. *Perception*, *7*, 247–260.