ACCESS: CONTEMPORARY ISSUES IN EDUCATION
1985, VOL. 4, NO. 1, 33–39

The effect of closed-loop causality

Felicity Haynes

Department of Education, University of Western Australia, Nedlands WA, Australia.

ABSTRACT
Underlying any analysis of classroom behaviour is some notion of causality, usually tacit and unconscious, and, more relevantly to this paper, usually linear, discrete and sequential. But underlying the notion of a linear causal chain there appears to be the assumption that external stimuli control the behaviour of organisms in a way analogous to the way in which internal forces acting on masses control the movements of physical bodies. Thus sensory stimulus and motor responses are believed to have as distinct physical existence as two billiard balls in motion. In this paper I intend to examine the consequences of replacing the stimulus-response model of causal explanations with a closed-loop model of causation which can include awareness and purpose.

> Electric circuiting is orientalizing the West. The contained, the distinct, the separate - our Western legacy - are being replaced by the flowing, the unified, the fused.
>
> M. McLuhan, The Medium is the Message.

Underlying any analysis of classroom behaviour is some notion of causality, usually tacit and unconscious, and, more relevantly to this paper, usually linear, discrete and sequential. Imagine if you will, the following scenario. It is a hot summer afternoon and the class is bored. Henry tilts his chair back onto John's desk. John, interpreting this as a hostile act, prods him in the back. Henry, who dislikes John for his consistently higher grades, retaliates by swatting John hard on the hand with his ruler, drawing blood. The teacher looks up and his glance is enough to make each boy settle back uncomfortably into his seat.

Prima facie, the scene is bound together by a chain of events like the following:

1. Hot weather caused general restlessness in the classroom.
2. Henry's restlessness caused him to tilt his chair back.
3. The bump from Henry's chair caused John to prod him in the back.
4. The prod caused Henry to react aggressively with a ruler.
5. The blow of the ruler caused injury to John's hand.
6. Events 2 to 5 caused the teacher to glance at the boys.
7. The teacher's glance caused the distraction to cease.

What, if anything is common to the notion of cause in each of these statements? At first appearance the fifth event stands out as the only non-behavioural one. It is subsumed under a general physiological law which claims that flesh subjected to a heavy blow will either bruise or bleed. Yet the first event is similarly physiological. It is generally known that extremely hot weather can cause pressure on the autonomous nervous system which can lead to restless behaviour. The difference is that an intelligent human, aware that heat can make one fidgety, can make an extra effort to concentrate or even take steps to cool himself thus breaking the causal necessity of a merely physical causal law.

The five remaining events involve human awareness to a greater extent. But underlying the linear causal chain we have presented seems to be an assumption that external stimuli control the behaviour of organisms in a way analogous to the way in which internal forces acting on masses control the movements of physical bodies. Thus sensory stimulus and motor responses are believed to have as distinct physical existence as two billiard balls in motion. In this paper I intend to examine the consequences of replacing the stimulus-response model of causal explanations with a closed-loop model of causation which can include awareness and purpose.

As far back as 1896 John Dewey was pointing out the limitations of the sensory-motor reflex arc concept in psychology. He claimed (1896, p. 360) that while it was an undoubted fact about human agency that movement as response follows sensation as stimulus, that such distinctions were not of fixed existence but distinctions of flexible function, that the cause had to be seen as cause before effect was possible, that one and the same occurrence could be either cause or effect, or stimulus or response depending on the shift of interest. The mistake Dewey' s psychology colleagues were making was to think in terms of an arc rather than a circuit, a circuit which was "more truly termed organic than reflex because the motor response determines the stimulus just as truly as sensory stimulus determines the movement. Indeed, the movement is often only for the sake of determining the stimulus, of fixing what kind of a stimulus it is" (Dewey, 1896, p.363).

In 1946, Bateson and Ruesch (p. 176) were insisting on the need for the same kind of circuit - "The network of value partially determines the network of perception ... it is also evident that perception determines values. As we seek things, so we act. But equally the success or failure of our action will determine our later vision. It is evident too that much of the change in the patient's system of values seems subjectively, to that patient, to be a change in the way he perceives things. Action would seem to be the middle term in which perception and value meet."

In 1974, an engineer named Bill Powers was saying in a book provocatively called "Behaviour: the Control of Perception" "What an organism senses affects what it does, and what it does affects what it senses. Only the first half of the commonplace observation has been incorporated into most psychological concepts of nervous system organization. The effects of behaviour in altering subsequent stimuli, and even in directly causing stimulation have certainly been noticed, but there has as yet been no correct analysis of this in any fully developed psychological theory (Powers, 1974, p. 41).

One can incorporate the second half of that circuit in a sequential manner. To revert to our example, Henry's action in causing his chair to tilt back caused him to receive the stimulus of John's prodding. A stimulus causes a response via the organism and the cycle begins again. As Hebb (1964, p. 58) said, "any behavioural response to a single stimulation thus produces a sensory feedback which can act as the initiator of a second response, whose feedback initiates a third response and so on".

But this is not what Dewey or Powers want. They claim that the real organism behaves in a smoothly continuous manner, with responses and stimuli continuously changing and continually interacting. Cause and effect are so smoothly co-ordinated that they are continuous, and ''the end gets thoroughly organised into the means" (Dewey, 1896, p. 366).

What we are now looking at is a theory of which Dewey was ignorant and Powers very much aware; the cybernetic theory of behaviour which says that systems which have inputs and outputs and which exist in an environment that makes the inputs depend in part on the output at the same time that the outputs depend on the inputs result in behaviour that tends strongly (a) to bring the input of the system to a preferred state called the reference level and (b) to vary so as to cancel nearly all the effect of any independent disturbance acting on that same input. As a result two components of behaviour can be distinguished. One component serves to bring the input to the reference level. The other serves to cancel the effects of disturbances which would otherwise make the input deviate from the reference level. Responses are thus dependent on present and past stimuli in a way determined by the current organization of the nervous system and equally, stimuli depend on responses according to the current organization of the environment and the body in which the nervous system resides. When Maturana and Letvin (1959, 1960) investigated frog and pigeon vision, they found that perception could not be viewed as a grasping of an external reality, but rather as the specification of one, because no distinction was possible between perception and hallucination in the operation of the system as a closed network. Indeed, in *Biology and Cognition*, Maturana argued that it was. inadequate and misleading to speak of causal relations when speaking about the circular organization of living systems - inadequate because the notion of causality is a notion that pertains to the domain of descriptions and as such is relevant only in the metadomain in which the observer makes his commentary and is not operative in the phenomenal domain. It is misleading because it obscures the understanding of the dependency of the unity on the distinctive operation that specified it.

In another attempt at simplification Petrie (1974) uses the example of a central heating system controlled by a thermostat where the sensor function senses the temperature in the immediate vicinity of the thermostat, which signal representing the actual temperature is compared with the reference signal, 22 C or whatever the arbitrary setting on the thermostat is. If there is a large enough difference, an error signal is sent to the effector function - the furnace, which puts out warm air which will affect the temperature which the thermostat senses. The feedback system is controlling its inputs where these are a function of both disturbances and the environmental feedback of the effect of the system itself. Under the linear causal chain model disturbances intervening between cause and effect change the effects, whence the long debates about what is to count as necessary and sufficient conditions. In the control theory model relevant disturbances are opposed so that the effect (the sensory stimulus staying near the reference level) stays the same. Note that for Petrie's thermostat model the colour scheme of the environment and the number of people in it are registered as stimuli or cause only insofar as they affect temperature, that is, cause a response in terms of the reference signal.

In these biological and mechanical examples the reference signals are a function of the organism either by genetic inheritance or an externally imposed restraint. They have no teleological function, or if so, only in an Aristotelian sense. In human interaction, values, language games, our forms of life, can set individual reference signals for different purposes, and moreover can be self-regulated or voluntarily changed. An example is the childhood game of hot and cold, where the subject's spatial relation to the hidden object causes the other players to say "hot" or "cold" and the subject's perception of these words is the basis of his actions that change his relationship to the object. We must assume a circuit that seems to be at the same time more adaptive and autonomous than machines or animals larking in self-consciousness, a point I shall have to argue for in more detail in a future paper. Before we discuss the more complicated human model, however, let us discuss a few other central features of the feedback system. The chains of causation common to organisms in environments, ecosystems, steam engines with governors, and societies, form circuits which are closed in the sense that causal interconnections can be traced around the circuit and back to whatever position was arbitrarily chosen as the starting point of the description. In such a circuit, events at any position in the circuit may be expected to have effects at all positions in the circuit at later times.

The feedback on such a circuit can be of two kinds, positive or negative. All evolving systems contain subsystems which are potentially regenerative, that is, they involve incremental change and would go into exponential runaway if uncorrected. On a cold morning a weak battery is unable to start the engine of a car at once and the continued use of the starter weakens the battery even further. The arms race escalates to a level which defeats its. original purpose to keep the peace. An organism's discomfort such as an itch activates a positive feedback loop to increase a scratching behaviour which may in fact increase the itch. In our classroom behaviour, Henry' s tipping his chair back may well have been an arbitrary. effort to alleviate discomfort caused by the heat, but if John had seen Henry as a jealous boy, John's prod could have been a device to test for hostile intent, in this case a self-fulfilling prophecy, one which reinforces the original perception in as Laing-ian a knot as an alcoholic who goes to do some controlled drinking in order to discover for himself that he has no control.

The regenerative potentialities of such subsystems are typically kept in check by various sorts of governing loops to achieve a steady state. They are homeostatic in that the effects of small changes of input will be negated and the steady state maintained by reversible adjustment. But a constancy of some variable is maintained by changing other variables. For instance, bicycles do not ride themselves. They require someone to ride them and convert a tendency to fall over into an action that will counteract this tendency - like a tilting of weight or a movement of the handlebars. In our example, the teacher acts literally as a governor to prevent the boys 9 hostile interactions from oscillating out of control. The sooner the teacher reacts to disturbances the less overt his control will be, but it is nonetheless effective. This is negative feedback.

Petrie applied his feedback model to a classroom situation to show the inadequacy of simply measuring the teacher's behaviour as if it were a stimulus causing student response, because there was no way of deciding on the basis of behaviour alone what caused the teacher sometimes to blame, sometimes to praise a student for what seemed identical behaviour. What appears to be contradictory behaviour could be seen as consistent in terms of a fixed reference signal which guides the teacher to try variant behaviours in order to match incoming perceptual signals with his reference signal, a reference signal which may say praise is only appropriate for students with a certain low level of self-esteem. So the teacher's behaviour can be said to cause quieter behaviour or higher achievement in only one respect. The teacher's reference signal (and of course, this must include an appraisal of the level of self-esteem each student has) has much greater explanatory force.

It is for this reason that we can begin to explain the sad lack of any necessary connection between the teacher's intention and the effect on the students. For, in my initial example, in order to see the teacher's glance as an inhibiting stimulus, the student's reference signal for disruptive behaviour must be set at approximately the same level as the teacher's. If it is not, he will not register his disapproval as such and it will not be effective. The problem of coupling self-corrective systems, especially conscious ones, is discussed at length by Bateson (1972, p. 446-453). Importantly, the boy's reference signal may make his disrupt behaviour desirable in terms of getting Susie' s approval, and Susie's delighted smile at his aggressive behaviour may cancel out the teacher's disapproving glare.

It would clearly be desirable to be able to change the boy's reference signal to one more compatible with the teacher's. This is a problem Petrie does not consider, and Powers is fairly pessimistic about the ability of any human being to control another. If the closed loop system were really closed, then there would be no way of changing the perceptions that counted and one's monitoring systems, the reference signals, would be as immutable as the innate governor that restricts the frog's visionary stimuli to moving objects in curved flight (Maturana etc., 1960). Powers says that we can make no deliberate effort to change our own reference signals without disastrous conflict and consequent pathology and to that extent we are literally governed by our reference signals. This is to make a very Kuhnian point that our conscious experience always has a point of

view which is partly determined by the nature of our learned perceptual functions involved and partly determined by built-in experience-independent criteria about which Piaget has had much to say. We cannot change our paradigm rationally, nor can we consciously change our reference signal. How then is learning possible?

Powers had never wanted to claim that man's behaviour is as mechanistic as a thermostat. This would pervert the original intention underlying his introduction of closed-loop causality, which was to bring to our attention the forces of intention, choice; freedom, reasons for acting, into behaviour. How can the system, if it is closed, avoid making the inbuilt adaptiveness, purposiveness, goal-seeking attributes of a cybernetic system either predetermined or arbitrary? If at first Powers' model seems to be pointing to some teleonomy, a first cause and final state, that may be because our reference signals constrain us to read him in terms of our well-rehearsed linear causality, and some primordial notion of external causes as physically real.

Powers' system remains open in a quasi-steady state, open like our scientific theories to gradual modification and evolution which we can modify only by stepping aside to some wider or alternative system. Powers' reference signals begin with genetically acquired neuronal systems but he claims that by a boot-strapping operation we can build up a hierarchy of eight, nine, perhaps ten, levels of organization of external stimuli, though he admits to a great deal of confusion as to what constitutes the level from which he is currently organising his theory. As Hofstadter suggests, logically we can never understand our own thought processes, though we can continually construct and reconstruct the level beneath the one we are operating from.

To relate this to the causal model we have been elaborating, we cannot explain the ultimate cause because to see it as a cause it must fit into a theory or loop which we have constructed ourselves and there will be some referencing signal of which we are as yet unaware influencing our description of it as a cause. The closed loop model argues against the physical existence of a cause qua cause, especially when it is operating to retain equilibrium. On the negative feedback model, the letter that one did not write can cause an angry reply. On the closed loop model effects are brought about by disturbances by discrepancies, by differences between the desired states-of-affairs and perceived actualities and such differences are themselves intangible.

So what is the current status of a cause? Relevant here is Hanson's perception (1969, Chapter 18) that causes are necessarily connected to their effects not because the world is held together by a kind of cosmic glue, but because the very notions of cause and effect are intelligible only against the background of a comprehensive theory of some sort. As scientists like Heisenberg, Born, Eddington have admitted, our scientific theories are systematized representations of the world only, flexible and adaptive *gestalten* in their own rights. *Pace* the behaviourist, there is nothing in the world we can kick that goes by the name of a cause. Cause functions as a proposition within a context and to see this allows us to include causal statements about purpose and reasons. Causality, like reason, is an instrument to bring order into our representations of experience. What is perceived as a disturbance to be counteracted is perceived as such against an individual's reference signal, whether that reference signal is an auto nomic nervous response, an awareness of pure form, a moral principle, a descriptive categorization. What is selected to count as a cause is anything which helps to understand a behaviour within a certain context of enquiry.

And yet, as Bateson had suggested, the relation between the real world and our theories is maintained by the fact that we act in a physical world. Unfortunately we have contact with the reference signals of other human beings only as they are manifested in external behaviour, whence the temptation of the behaviourists to speak as if that were all that mattered. There are many explanatory loops into which we can fit our causal notion. In our example, we could include the strong parental expectation that Henry must succeed at school and his resultant feelings of frustrated inadequacy to meet their perceived reference signal. To explain why Henry over-reacted this time and not at other times would perhaps include the closed-loop effect of the heat on the autonomic nervous-system within the higher social conflict loop, and such a loop would require a

different closed loop than the explanation as to why Henry hit John on the hand rather than the head, a loop which may include the reference signals controlling desired classroom behaviour that also influenced the final settling down.

This is beginning to sound as messy as many of the current structuralist theories. But it is no worse than the current arguments about causality that are still appearing in *Mind*. It is similar to Gasking's recipe model in which causes are seen as an intellectual tool to name certain conditions. within the control of agents by which they can bring about or prevent other conditions. "One says 'A causes B' in cases where one could produce an event or state of the A sort as a means to producing one of the B sort." This makes possible not only the manipulation of certain objects, but human purposes become part of the definition of cause. It is also a flexible explanatory construct in the sense that the answer as to why John hit Henry depends on why you are asking why, and whether you could use the same answer to prevent it happening again.

Even though Gasking admits that cause can occur simultaneously with effect, his examples treat cause and effect as an orderly march from stimulus-object to sensory receptors and from muscle tension to the eventual behavioural event. But this need not be because cause is actually a matter of external existence or sequential progression. The linear and discrete constraints of language make one frame a verbal account of cause in such a way as to make events appear isomorphic with the language in which they are expressed. Taking the whole circuit into account and including purpose cannot prevent one from describing it as "O is desirable; B leads to C; C leads to D; so D can be achieved by way of Band C" even if the achievement of Din fact is continuous, co-ordinated and unconscious.

There seems no way out of this impasse, but it need not bother us if we are aware that the linearity of cause is due to the limits of our language. Bateson (1972, p. 450) suggests that because the past cannot logically encompass the whole, the cybernetic nature of self and the world must be imperceptible to consciousness which forms only a part of the human system. Thus our conscious sampling of data will not disclose whole circuits but only arcs of circuits, cut off from their matrix by our selective attention. Specifically, the attempt to achieve a change in a given variable, located either in self or in environment, is likely to be undertaken without awareness of the homeostatic network surrounding that variable.

Ironically, that means that despite the partial success of the closed-loop theory in blurring the dichotomy between cause and effect there seems no way of doing away with a sequential and discrete description of them.as two separate components of the same loop, no matter how complex and continuous the circuit becomes. I have been guilty of the same implication in this paper when I state that the reference signals control behaviour, as if the behaviour was not also controlling the reference signal. The only saving grace is that the rules of language protect us as well as limit us. Cause X can explain effect Y, but no matter how close these can actually come together on a continuum, or how far one is influenced by the other, in our language game we have stipulated that the effect cannot explain the cause.

The largest problem about claiming that the notion of causality is largely constrained by our language is that it seems to make causes inaccessible to scientific enquiry. But in fact that is no less true of causes than it is of any scientific constructs. While causes qua causes exist only as a logical sequence in descriptions and theories, they can be inferred from the real world by testing hypotheses. This is so even in the case of non-existent letters, purposes and invisible reference signals. The crucial test adapts a property used by servo-mechanism engineers to measure properties of control systems, namely to apply a known disturbance to the quantity thought to be controlled and observe in detail the subsequent behaviour of the quantity under the influence of continuing steady disturbance and the behaving system's output. Powers' rubber band experiment (1974, p. 241-4) illustrates feedback causation quite simply. A subject (S) and experimenter (E) each put a forefinger in an end of a pair of rubber hands looped together and hold the bands slightly stretched over a table-top. Sis told to keep the knot stationary over some inconspicuous mark on

the table-top. E can disturb the position of the knot by pulling back or relaxing his pull on his end of the rubber band pair: S maintains the knot where he wants it by similar means. If E draws slowly back on his end of the rubber, Swill do likewise, and the knot will remain quite stationary. If E swings his end slowly from side to side, Swill swing his end the opposite way, and the knot will still not move. From the behaviouristic point of view, E's finger movements constitute the stimulus, and S's the response. If E can reach any conclusion, it would most likely be that the subject's response is generally (statistically) opposite to the stimulus in direction, especially if E's movements are rapid or jerky. But unless E happens to notice that the knot stays still most of the time, he will miss the crucial features of the situation - the purpose of S's every movement. In ordinary behavioural situations, the "rubber bands" are hidden or invisible, and the knot - the controlled quantity - is far from obvious. All that is obvious is the relationship between the disturbance and the subject's output. But the obvious is clearly not the only causal factor, nor indeed the most important one. We can infer causes from purposes, unconscious body processes, physical objects in much the same way that physicists can infer quarks from controlled experimentation.

What then arc the effects of the closed-loop model of causality? Its principal advantage is that it makes us aware of a broader context within which causal chains are usually presented. It shows that causal claims, like all epistemological claims are theory-dependent and because of that need necessarily to be presented as discrete and sequential, though the process they describe need not be. There may well be among humans a hierarchy of causal explanations, including overriding ethical principles, tacit agreements, which are isomorphic with physical or physiological explanations. Thus an agent's actions can be described as caused by something at one level while at the same time allowing the agent to have chosen to operate at that level. The closed loop model can accommodate the criticisms that Dewey made of the reflex arc model, by including teleological explanations, continuity and co-ordination within its scope and still allowing for the behaviourist model to be seen as part of the picture. Though we remain restricted to describing cause and effect as though they were discrete entities the discrete quality comes from the nature of abstracted referential representation and not necessarily from the cause itself. It should be clear to classroom observers that the closed-loop theory of causality at least legitimizes causal inferences from invisible reference signals and that observed stimuli need only be a side-effect of deeper teleological causes, such as desires, emotions, volition, intention and even moral sanctions.

References

Bateson, Gregory (1972) *Steps Towards an Ecology of Mind*, Part V, Chandler, San Francisco.

Dewey, John (1896) "The reflex arc concept in Psychology", *The Psychological Review*, Vol. III, No. 4, July, pp. 357ff.

Gasking, Douglas (1955) "Causation and Recipes", *Mind*, Vol. 64, No. 256, October, pp. 479-487.

Hanson, N.R. (1969) *Perception and Discovery*, Freeman, Cooper and Co.

Hebb, 0.0. (1964) *A Textbook of Psychology*, Philadelphia: Saunders.

Hofstaoter, Douglas (1969) *Godel, Escher, Bach: An Eternal Golden Braid*. Sussex: Harvester Press.

Maturana, H.R., Lettvin, J.R., McCulloch, W.S. & Pitts, W.H. (1960) "Anatomy physiology of vision in the frog" (Rana pipiens), *Journal of General Physiology,* Vol. 43, No. 6, Part 2, pp. 129-175. and

Maturana, et al. (1959) "What the frog's eye tells the frog's brain", *Proceedings of the IRE*, Vol. 47, No. 11, pp. 1940-1959.

Petrie, Hugh (1974) "Action, Perception and Education", *Educational Theory*, Vol. 24, No. 2, pp. 33-45.

Petrie, Hugh (1981) *The Dilemma of Enquiry and learning*. University of Chicago Press.

Powers, W.T. (1974) *Behaviour: The Control of Behaviour*, Chicago: Aldine Press.

Ruesch, J. & Bateson, G. (1951) *Communication: The Social Matrix of Psychiatry*, New York: W.W. Norton.