Control theory vs. Reinforcement theory

This thread continues where Research_PCT.pdf ends. The dialog on rat experiments continues in ShockingExperiments.pdf.

Unedited posts from archives of CSG-L (see INTROCSG.NET):

The file StarTrek.pdf holds an earlier discusion of operant conditioning, which is part of Reinforcement Theory.

[The term EAB stands for Experimental Analysis of Behavior -- the study of conditions which influence operant behavior]

Date: Tue, 6 Jun 1995 02:44:18 -0600
Subject: reinf theory;

[From Bill Powers (950606.0000 MDT)]

>[Bruce Abbott (950605.1210 EST)]

> One difficulty is that there _is_ such a thing as reinforcement -- by definition. So when you tell a reinforcement theorist that there "ain't no such thing," you're bound to be met with a bit of skepticism.

You just have to separate what is observed from what is imagined. What is observed is that animals will produce actions that produce food, water, etc., and that given practice they will do this more and more reliably. This way of putting it gives us a complete description of what is observable, without ever mentioning "reinforcement" as a thing or a process.

We can easily verify the "forward" link; we can show that the consequences of behavior follow lawfully from the behavior. The laws are simply those of physical science and engineering. We can also easily verify that as the behavior gradually changes, the consequences change, too. On these observations there will be no disagreement between behaviorists and PCTers.

What is imagined is that the food, water, etc. are causing the behavior of the organism to change as we see it changing. This "reverse link," however, is not observable or verifiable. We find that the only possible verification reduces to a restatement of the observable "forward" relationships.

The transition to PCT comes when we propose a model for the reverse relationships, the ones connecting the consequence of behavior back to the behavior via the organism. Doing this using correct mathematics leads inevitably to the equations of a closed-loop system, and the transition to a systems analysis of behavior rather than a cause-effect analysis. We find that the so-called reinforcer is just a variable in a closed loop of relationships. We learn to distinguish performance (operation of a closed loop with fixed parameters) from learning or adaptation (processes that alter the parameters of a given closed-loop system).

That's one line of argument.

> If I food-deprive a rat for a few hours and then give the rat the opportunity to earn a bit of food by pressing a lever, the rat learns to press the lever. The frequency of lever-pressing will increase over time and will be maintained at some higher level so long as the rat remains hungry and the contingency between lever-pressing and pellet delivery remains in effect.

What is important is that the food-deprived rat will learn to do whatever is required to reduce the food deficit. By stopping at the lever-pressing, you are only noticing one point in a closed loop. You have to carry the analysis all the way around the loop, wherever you started. If you started with a reduction in food intake, you have to follow the process all the way around to an effect on food intake; the effect is to act against the deficit. If you do the equivalent of pressing the lever now and then, independently of the rat, you can predict what will happen by tracing events all the way around to the lever-depressions again: assuming that the system is inside the normal control range, the rat will reduce its own lever-pressing, acting against the externally-induced increase in rate of depression of the lever. Disturbances applied anywhere in the loop are resisted.

> The observed increase in responding is _by definition_ reinforcement.

But notice that if you artificially elevate the rate of delivery of food-pellets, the most likely effect is a _reduction_ in the rat's rate of lever-pressing, because (going the rest of the way around the loop), the change in food delivery rate is opposed by the change in the behavior rate. This assumes, of course, that the system is inside the normal control range, and not at an extreme of deprivation where the relationships reverse.

You have still never explicitly commented on this problem: that you get the assumed relationship between increases of behavior and increases in reinforcement only for the most extreme degrees of deprivation. Is this something that behaviorists have simply agreed to keep quiet about? Some of them certainly haven't agreed to this.

> Thus the argument is not whether there is or is not a phenomenon which has been labeled "reinforcement," but how this phenomenon is to be explained.

But in fact, over most of the possible range of schedules, the relationship observed is that an increase in reinforcement rate leads to a decrease in behavior rate. So I would say that there is a serious problem with assuming that there is no argument over the existence of the phenomenon. I can make a good case that the phenomenon is largely an artifact of experimental conditions. Standard experimental conditions have been adjusted until the theoretically expected relationship is observed.

> Reinforcement theory holds that certain sensory consequences of responding cause changes to occur in the nervous system that make it more likely that the response will occur again. PCT holds that responding increases because error between a controlled perception and its reference level generates output that tends to reduce the error; given the environmental feedback function, reduction of error requires an increase in (lever-pressing) output. Reinforcement theory makes reinforcement the central explanatory principle for behavior change; PCT makes it a side-effect of control. The argument is not about the objective phenomenon of reinforcement but its theoretical significance.

But PCT predicts that any externally-induced change in the rate of reinforcement will result in an _opposite_ change in the rate of behavior, while reinforcement theory predicts that the change in rate of behavior will be _in the same direction_ as the change in rate of reinforcement. This is a very clear and unequivocal contradiction.

We can find copious data in the published literature to support both predictions. When we examine the data to see how this can possibly be true, we find that the data supporting each side were taken under very different experimental conditions. The PCT prediction is supported clearly by data taken under conditions of mild to large degrees of deprivation, but generally those degrees of deprivation that could still allow life to continue if extended indefinitely. The reinforcement-theory prediction is supported under conditions of extreme deprivation, under which the animal could not survive.

The data I know about show that the transition (for food reinforcement) occurs at a rate about 1/3 of the normal free- feeding food intake. A human being who consumes about 3000 calories per day would just be on the boundary between the two conditions on a diet providing 1000 calories per day. At this level of input, reinforcement theorists would say that "satiation" has occurred, because further increases in food intake would fail to increase the behavior rate (under the same kinds of ratio schedules used to take the rat data). To get a clear demonstration of the positive relationship between reinforcement and behavior, the food intake would have to be in the range of perhaps 100 to 800 calories per day. In other words, human subjects would be starving to death.

Starvation is warded off in operant conditioning experiments by adjusting the food intake between experiments. Reducing a rat's weight to a stable 80 percent of free-feeding weight requires, I would guess, lowering average daily food intake by less than 10 percent. In obesity experiments, it has been found that rats with hypothalamic lesions, who may stabilize their free-feeding weight at 300% of preoperative weight, eat only about 5% to 10% more than before, after they reach equilibrium.

Therefore a 70% reduction in food reinforcement rate, if sustained, would undoubtedly be lethal. And _at least_ this degree of deprivation during an experiment is required in order to show a clear positive relationship between behavior and reinforcement.

So the very existence of reinforcement as a phenomenon is in question.

Best to all, Bill P.

Date: Tue, 6 Jun 1995 14:09:32 -0500
Subject: Reinforcement: Keeping Quiet

[From Bruce Abbott (950606.1400 EST)] >Bill Powers (950606.0000 MDT) --

>>Bruce Abbott (950605.1210 EST)

>> The observed increase in responding is _by definition_ reinforcement.

> But notice that if you artificially elevate the rate of delivery of food-pellets, the most likely effect is a _reduction_ in the rat's rate of lever-pressing, because (going the rest of the way around the loop), the change in food delivery rate is opposed by the change in the behavior rate. This assumes, of course, that the system is inside the normal control range, and not at an extreme of deprivation where the relationships reverse.

> You have still never explicitly commented on this problem: that you get the assumed relationship between increases of behavior and increases in reinforcement only for the most extreme degrees of deprivation. Is this something that behaviorists have simply agreed to keep quiet about? Some of them certainly haven't agreed to this.

I believe that you have drawn an incorrect conclusion from one graph showing what happens to asymptotically maintained response rates as the ratio requirement of a ratio schedule is varied. Perhaps a clearer picture might emerge from examining data on runway performance. Rats are food-deprived to some criterion and then trained to run from a start box, down a straight-alley, and into a goal box which contains a certain amount of standard laboratory rat chow. What is measured is the speed of running. Running speed will vary directly with (a) the amount of food provided in the goal box and (b) the level of deprivation (so long as the level is not so high as to begin to incapacitate the rat).

The rats are run for only one or at most only a few trials per day so as to prevent much of a satiation effect. Both of the manipulations (deprivation and amount of food) may be viewed as affecting the attractiveness of the goal. For a rat not AT the goal, the experienced error for not being at the goal will be proportional to the goal attractiveness. Larger errors will lead to greater output (faster running).

In this description I have oversimplified, especially as there are multiple control systems at work governing perceptions at several levels and I have improperly collapsed these into one, but I think I've communicated the basic idea: larger reward would be expected to produce faster running if you properly analyze the situation from a control-system perspective (remember, each visit to the goal box is having little or no effect on the error in the upper-level nutritional control system even with relatively large reward; the experimental procedure essentially opens the loop on this system).

An equivalent manipulation to changing the ratio would be to increase the distance the rat has to run to reach the goal. Now it must run for a longer time and use more energy to reach its perceptual reference state of being in the goal box, eating the food. If the trip is not too taxing the extra effort may have little effect on running speed, and experimenter will have gotten more distance out of the rat for the same reward. The curve will begin to reverse only when the effort expended begins to seriously detract from the attractiveness of the goal. If the analysis takes into account the "reinforcement" received at the goal AND the "punishment" received in terms of fatigue, the results appear to be quite compatible with a reinforcement analysis rather than something that one needs to "keep quiet about." Staddon certainly didn't see anything problematic. And these results have absolutely nothing to do with operating at the extreme end of the deprivation scale. The rat's control system is operating within its normal range, but a key component of that (hierarchical) system is operating with its feedback loop open. Larger errors lead to larger output, until error in another control system (fatigue) begins to have an increasingly large opposite effect on the output.

I realize that in the above description I have not done full justice to either the reinforcement or PCT analysis of the experimental results as I have not yet had a chance to think the problem through in detail, but I do believe I have captured at least the essence of the explanation.

Regards, Bruce

Date: Tue, 6 Jun 1995 17:18:03 -0600
Subject: reinforcement problem

[From Bill Powers (950606.1440 MDT)]

Bruce Abbott (950606.1400 EST) --

> I believe that you have drawn an incorrect conclusion from one graph showing what happens to asymptotically maintained response rates as the ratio requirement of a ratio schedule is varied.

Oh? In what way was it incorrect? There are many more such graphs, including some obtained by Timberlake showing the negative relationship between reinforcement rate and behavior rate over a range of ratios of 5000:1, with NO region in which the positive relationship was observed. In these experiments the animals obtained all food or water from the experimental apparatus; the experiments ran continuously. In obesity experiments, mentioned in BCP, it was found that adding reinforcers arbitrarily caused an immediate reduction, even a cessation, in behavior, and ceasing to add them immediately restored the former rate of behavior. There is even a principle I've heard of: "noncontingent reward decreases behavior."

> Perhaps a clearer picture might emerge from examining data on runway performance. Rats are food-deprived to some criterion and then trained to run from a start box, down a straight- alley, and into a goal box which contains a certain amount of standard laboratory rat chow. What is measured is the speed of running.

The picture isn't "clearer;" it's just more like what you would expect from reinforcement theory. In this sort of experiment, speed of running has very little effect on amount of reinforcement received, and since as you say only one or a few trials per day were run, there is no possibility of measuring the slope of the behavior-reinforcement curve. Most of the time is spend in another cage under another schedule, which is not mentioned. The experimental run is just a blip in the background conditions.

I wonder why only a few runs per day were used. Could the reason be that if the rat spent a few hours running the T-maze, the expected relationship would no longer be seen? Presumably, if you put a lot of food in the goal-box on each run, the rat would slow down, and come to an asymptotic speed that would be slower as the amount of food increased -- at least above some critical amount of food. With enough food in the goal-box, there would be trials on which the rat didn't bother to run at all. This is what is interpreted as "satiation," which is to be avoided as it produces results not consistent with the theoretical prediction. And if the amount of food was decreased, the rat would speed up, trying to get more food. However, if the amount of food was made small enough, we would start to see the other relationship: increasing the amount of food would result in faster running, decreasing it in slower running.

That's my prediction, based on the other experiments. Any indication that it is wrong?

> .. larger reward would be expected to produce faster running if you properly analyze the situation from a control-system perspective (remember, each visit to the goal box is having little or no effect on the error in the upper-level nutritional control system even with relatively large reward; the experimental procedure essentially opens the loop on this system).

The only problem is that in order to make the results come out right, you have to postulate something unobservable: changes in the "attractiveness" of the goal with a sign chosen to be just right to explain the results. All we actually observe is that the rat, fresh from the 23 hours of training in the other cage, will strive harder to get food when it sees more food available, and when it has been more deprived, up to a point.

If a larger reward is expected to produce faster running, what happens when the reward exceeds the amount of food that the rat normally eats in a day? Will the running speed go off the upper end of the scale? Or will it slow to a stroll, a saunter, a causal exploratory run with much sniffing here and there and eventual leisure arrival for a nibble at the food in the goal-box?

I believe that the conditions in this experiment have been carefully adjusted until the relationship predicted by reinforcement theory was observed.

> Staddon certainly didn't see anything problematic.

That's right. He didn't seem to notice anything funny about the curves. He didn't comment on their significance relative to basic reinforcement theory. Maybe he just missed seeing the problem, or maybe he decided he wasn't going to touch this one with a ten-foot pole.

The main thing I get from your post is that if my interpretation of the data is right, there is definitely a problem for reinforcement theory and you would feel a considerable urge to make it go away.

Best to all, Bill P.

Date: Tue, 6 Jun 1995 19:09:28 -0700
Subject: Reinforcement: The Control of Perception

[From Rick Marken (950606.1900)]

Bruce Abbott (950606.1400 EST) --

> I realize that in the above description I have not done full justice to either the reinforcement or PCT analysis of the experimental results as I have not yet had a chance to think the problem through in detail, but I do believe I have captured at least the essence of the explanation.

Actually, your description of a reinforcement analysis of behavior explains quite a bit. When you know what to look for, controlled perceptions are really quite obvious. Of course, Bill Powers (950606.1440 MDT) picked up on it too:

> The main thing I get from your post is that if my interpretation of the data is right, there is definitely a problem for reinforcement theory and you would feel a considerable urge to make it go away.

Best Rick

Date: Wed, 7 Jun 1995 12:56:05 -0500
Subject: Caught!

[From Bruce Abbott (950607.1245 EST)]

>Bill Powers (950606.1440 MDT) --

> The main thing I get from your post is that if my interpretation of the data is right, there is definitely a problem for reinforcement theory and you would feel a considerable urge to make it go away.

Alternatively, if my interpretation of your interpretation is correct, there is definitely a problem with your understanding of reinforcement theory and I would feel a considerable urge to make THAT go away. (:->

>Rick Marken (950606.1900)] --

>>Bruce Abbott (950606.1400 EST)

>> I realize that in the above description I have not done full justice to either the reinforcement or PCT analysis of the experimental results as I have not yet had a chance to think the problem through in detail, but I do believe I have captured at least the essence of the explanation.

> Actually, your description of a reinforcement analysis of behavior explains quite a bit. When you know what to look for, controlled perceptions are really quite obvious. Of course, Bill Powers (950606.1440 MDT) picked up on it too:

>> The main thing I get from your post is that if my interpretation of the data is right, there is definitely a problem for reinforcement theory and you would feel a considerable urge to make it go away.

For some reason, I'm getting an echo in some of the csg-l posts I have been receiving. Gary, perhaps you could check on it, it's wasting bandwidth.

O.K., O.K., you caught me. All this time I've been part of a top secret, high-level plot to undermine PCT and reestablish traditional reinforcement theory as the "top dog" in the field of learning and behavior (after we do the same to cognitive theory, of course). It's been my job to infiltrate CSG-L and discover, through carefully designed, probing attacks, the any weaknesses that could be used, fairly or unfairly, to discredit perceptual control theory and all those who espouse it. Yeah, yeah, that's the ticket! Yeah, I'm a double agent, an' I get secret messages from B.F.S., who is, ah, not really dead, see? Yeah, he's just lyin' low until the right moment, ready to take over an' turn the world into one gigantic operant chamber, see? An the only thing in our way is PCT (an' maybe them cognitive guys), yeah, an' it's TOO LATE to stop us, we got guys all over CSG-L. Why do you think you get so many arguments? In fact, we're ALL in on it, everyone except Tom Bourbon, yeah, that's it, so we had RID of him. You're FINISHED, see?

Rick Marken:

> Actually, your description of a reinforcement analysis of behavior explains quite a bit. When you know what to look for, controlled perceptions are really quite obvious.

Bill Powers:

> That's how the test for the controlled variable works: you can't prove beyond doubt that a particular variable is under control, but you can very quickly show that it's not under control.

Hmmmm. Rick, you don't have a CLUE as to what my motives are.

Regards, Burrhus, er, Bruce

Date: Wed, 7 Jun 1995 23:14:58 -0600
Subject: PCT & reinforcement

[From Bill Powers (950607.2010 MDT)]

Bruce Abbott (950607.1245 EST) --

> O.K., O.K., you caught me. All this time I've been part of a top secret, high-level plot to undermine PCT and reestablish traditional reinforcement theory as the "top dog" in the field of learning and behavior

I don't see that as the problem. What I see is the Necker Cube problem, very difficult to overcome. Consider this:

> I believe that you have drawn an incorrect conclusion from one graph showing what happens to asymptotically maintained response rates as the ratio requirement of a ratio schedule is varied.

This way of saying it makes the assumption that it is the response rate that is being maintained by the reinforcement rate. However, when for example the schedule is changed from FR1 to FR2, the first thing that happens is that the reinforcement rate drops in half. It will remain at that level unless the behavior rate changes. As the behavior rate increases, the reinforcement rate will increase, because it is dependent on the behavior rate. What you end up with is a reinforcement rate that is somewhat less than before, and a behavior rate that is almost twice as high as before.

Clearly, it is the reinforcement rate that is being maintained by the behavior rate. The direction of the dependency is easy to verify in a number of ways, a simple way being to examine in detail what happens immediately after a change in the schedule. Another simple way is to examine the apparatus, which will show exactly how variations in behavior cause variations in reinforcement rate. On a ratio schedule, it is impossible for the reinforcement rate to be anything but the behavior rate divided by the ratio.

Best, Bill P.

Date: Thu, 8 Jun 1995 13:21:02 -0700
Subject: Alternative theories and alternative phenomena

[From Rick Marken (950608.1320)]

Bruce Abbott (950605.1210 EST) --

> Reinforcement theory makes reinforcement the central explanatory principle for behavior change; PCT makes it a side-effect of control. The argument is not about the objective phenomenon of reinforcement but its theoretical significance.

This statement suggests that the same phenomenon (behavior change) is explained by both reinforcement theory and PCT; the two theories just explain it in different ways: reinforcement theory says that behavior change results from the "strengthening" effect of reinforcement; PCT says that behavior change is a side effect of control.

I think Bruce is claiming that PCT is better (more accurate, simpler, etc) than reinforcement theory as a model of behavior change. In other words, reinforcement theory (like Ptolemaic theory) is basically OK but PCT (like Copernican theory) is much better. Bruce is, therefore, understandably puzzled by the response he is getting from those of us who presumably also believe that PCT is better than current theories of behavior. This puzzlement is obviously very frustrating for Bruce, as evidenced by:

> "O.K., O.K., you caught me. All this time I've been part of a top secret, high-level plot to undermine PCT and reestablish traditional reinforcement theory as the "top dog" in the field of learning and behavior..."
> Bruce Abbott(950607.1245 EST)

Clearly, Bruce sees himself as a friend of PCT who is being treated as an enemy. I think other friends of PCT have felt the same way. What's going on?

It think what's going on is a conflict between two very different views of PCT. One view (Bruce's) is that PCT is another theory of behavior -- like reinforcement theory, various cognitive theories, motivational theories, etc -- and is of interest to the extent that it can explain existing data _at least_ as well as the other theories can. This is the "alternative theory" view of PCT. The other view (mine) is that PCT is about the phenomenon of control and that the goal of PCT is to understand this phenomenon. According to this view, most existing data is irrelevant to understanding control. This is the "alternative phenomenon" view of PCT.

These two views can sometimes become somewhat conflated. For example, in his reply to Bruce's quote above, Bill Powers (950606.0000 MDT) says:

> But PCT predicts that any externally-induced change in the rate of reinforcement will result in an _opposite_ change in the rate of behavior, while reinforcement theory predicts that the change in rate of behavior will be _in the same direction_ as the change in rate of reinforcement. This is a very clear and unequivocal contradiction.

This gives the impression that Bill sees PCT as an alternative to reinforcement theory. In fact, Bill is comparing a _qualitative_ prediction of reinforcement theory to a quantitative prediction of a PCT model. In fact, there is no such thing as a reinforcement theory of control. A model that actually produces a change in rate of behavior (output) that is in the same direction as the change in rate of reinforcement (input) is a positive feedback model; it doesn't keep the input (or anything) under control.

Once you know that the rat in an operant chamber is controlling food input (you have identified the phenomenon of control), you know that only a control model (like PCT) can explain the phenomenon; a reinforcement is simply not an alternative.

In general, when know there is control, then you know there is no alternative to control theory.

Best Rick

Date: Thu, 8 Jun 1995 17:07:51 -0600
Subject: Two strategies

[From Bill Powers (950608.1600 MDT)]

Rick Marken (950608.1320) --

> This gives the impression that Bill sees PCT as an alternative to reinforcement theory. In fact, Bill is comparing a _qualitative_ prediction of reinforcement theory to a quantitative prediction of a PCT model.

Our strategy here is a little different. You are going directly for the idea of PCT as the theory of control, with other theories being about something else -- output production, mostly. I agree with this, as you know.

But another way to deal with a disturbance beside directly resisting its effects is to look for its cause and remove it, if possible. If the cause of the disturbance can be removed, less effort will be needed to resist it -- in fact, none.

In my discussion with Bruce about reinforcement theory, I'm trying to show that regardless of what PCT says, reinforcement theory doesn't fit all the facts of behavior and is straight-out incorrect even as a description of what happens under most real conditions. The other main line of argument is that we have many ways to show that reinforcement, as a variable, is a dependent variable, not an independent variable. I think we are solid on these points, and that they spell the end of reinforcement as an explanatory construct. Of course there will always be those who find complicated arguments to preserve the status quo, but we can't do anything about that. Some will get the point.

Best, Bill P.

Date: Thu, 8 Jun 1995 19:35:20 -0500
Subject: <No subject given>

[From Bruce Abbott (950608.1930 EST)]

>Rick Marken (950608.1320)] --

>>Bruce Abbott (950605.1210 EST)

>> Reinforcement theory makes reinforcement the central explanatory principle for behavior change; PCT makes it a side-effect of control. The argument is not about the objective phenomenon of reinforcement but its theoretical significance.

> This statement suggests that the same phenomenon (behavior change) is explained by both reinforcement theory and PCT; the two theories just explain it in different ways: reinforcement theory says that behavior change results from the "strengthening" effect of reinforcement; PCT says that behavior change is a side effect of control.

Yep.

> I think Bruce is claiming that PCT is better (more accurate, simpler, etc) than reinforcement theory as a model of behavior change. In other words, reinforcement theory (like Ptolemaic theory) is basically OK but PCT (like Copernican theory) is much better.

Nope.

I am claiming that current reinforcement provides a consistent and generally compelling framework for understanding behavior (which is why it has so many adherents) and is capable of handling the sort of data Bill P. asserts it cannot (the ratio data). This is not the same as saying that it is basically O.K., which seems to be your take on what I am saying. One can grant that Ptolemaic theory correctly describes the apparent motions of the planets through the heavens without implying that the theory is correct ("O.K.")

If the theory can handle such data then its adherents will see nothing surprising or contradictory in such data, nor will they feel the need to sweep such findings under the rug. This explanation eliminates the need to posit any conspiracy of silence.

Bill P. noted that I had avoided the issue about performance on ratio schedules when it came up earlier and I did not wish to appear to be ducking it. So I attempted to respond while realizing that I had not really thought the problem completely through. The result is that I did not really make a good case and I knew it, but I still believe such a case can be made, once I've given it more consideration.

> Bruce is, therefore, understandably puzzled by the response he is getting from those of us who presumably also believe that PCT is better than current theories of behavior. This puzzlement is obviously very frustrating for Bruce, as evidenced by:

>> "O.K., O.K., you caught me. All this time I've been part of a top secret, high-level plot to undermine PCT and reestablish traditional reinforcement theory as the "top dog" in the field of learning and behavior..."
Bruce Abbott(950607.1245 EST)

Well, it is frustrating when your motives (er, attempts to correct deviations of controlled perceptions from their reference values) are misunderstood, as in this case, especially when one of the people doing the misunderstanding claims to be able to read your mind (as in his mindreader program). But the paragraph from which the above quote was taken was intended to do something Rush Limbaugh likes to do: illustrate absurdity with absurdity. [No, I'm not a big Limbaugh fan, thank you.] After all, the above "admission" is also consistent with the "evidence," isn't it? And if it is, how can you claim to know what I'm "really" trying to do? Hey, it might even be true... (;->

I do agree that a part of the difficulty in challenging reinforcement theory is that many of its proponents do evaluate it only in qualitative terms. It is also the case that reinforcement theory is currently undergoing challenges and modifications in response to those challenges, so that it constitutes a moving target or, perhaps more accurately, a number of alternative views. For this reason a given finding may be fatal to one version but consistent with another.

> Clearly, Bruce sees himself as a friend of PCT who is being treated as an enemy.

Not as an enemy, but perhaps as someone whose viewpoint is sometimes misunderstood and criticized for the wrong reasons. I would hope that there are no "enemies" here, just people with sometimes differing opinions who are willing to argue them and to listen carefully to the other point of view. But yes, I do see myself as a friend of PCT.

> It think what's going on is a conflict between two very different views of PCT. One view (Bruce's) is that PCT is another theory of behavior -- like reinforcement theory, various cognitive theories, motivational theories, etc -- and is of interest to the extent that it can explain existing data _at least_ as well as the other theories can. This is the "alternative theory" view of PCT. The other view (mine) is that PCT is about the phenomenon of control and that the goal of PCT is to understand this phenomenon. According to this view, most existing data is irrelevant to understanding control. This is the "alternative phenomenon" view of PCT.

I do think of PCT as another theory of behavior, but not as one alternative among many equals. Rather, I think of it as the correct solution, at least in broad outline (the details remain to be worked out, remember). Behavior is what I'm interested in studying and understanding. I'm not interested in studying the "phenomenon of control" unless it helps me to understand behavior. It does, so I am. But the "phenomenon of control" is only a part of the total picture: there is the phenomenon of perception, the phenomenon of memory, the phenomenon of learning, the phenomenon of discrimination, and many others, all worthy of study in their own right.

> Once you know that the rat in an operant chamber is controlling food input (you have identified the phenomenon of control), you know that only a control model (like PCT) can explain the phenomenon; a reinforcement is simply not an alternative.

Well, I'm convinced of that, but I wish to convince others who currently see the reinforcement explanation as adequate, and to do so I need to understand how the PCT model of the universe accounts for the apparent motions of the planets, even though those motions are only a side-effect of not being at the center of the universe.

Regards, Bruce

Date: Thu, 8 Jun 1995 21:35:09 -0700
Subject: Let's see the reinforcement model, already

[From Rick Marken (950608.2130)]

Bill Powers (950608.1600 MDT) --

> In my discussion with Bruce about reinforcement theory, I'm trying to show that regardless of what PCT says, reinforcement theory doesn't fit all the facts of behavior and is straight-out incorrect even as a description of what happens under most real conditions.

I think it's a fine approach and you can count me in. As a matter of fact I think I see one of them "reinforcement theory fits the facts" fella's headin' up over the ridge now.

Bruce Abbott (950608.1930 EST)--

> I am claiming that current reinforcement [theory] provides a consistent and generally compelling framework for understanding behavior (which is why it has so many adherents) and is capable of handling the sort of data Bill P. asserts it cannot (the ratio data).

Well, I'd say it's time for you to show how reinforcement theory handles the ratio data, pardner. I'm talking about a workin' model, friend; no curve fitten'. So it won't do ya' no good ta draw =|;-)

> Bill P. noted that I had avoided the issue about performance on ratio schedules when it came up earlier... I did not really make a good case and I knew it, but I still believe such a case can be made, once I've given it more consideration.

You remind me of a Canadian fella' who used to keep sayin' the same thing about information theory =|;-)

> [No, I'm not a big Limbaugh fan, thank you.]

I can't believe anyone watches him at all. Have we completely lost our wits (Swift, Mencken, Twain, Vonnegut).

Bruce Abbott (950608.1750 EST)

> Your Honor, my client would be most interested if counsel would be willing to share his insights with the court.

I guessed what you were controlling for in another post and you basically agreed with my description. I said:

> I think Bruce is claiming that PCT is better (more accurate, simpler, etc) than reinforcement theory as a model of behavior change. In other words, reinforcement theory (like Ptolemaic theory) is basically OK but PCT (like Copernican theory) is much better.

You disagreed with the "OK" part. So I will change " reinforcement ... is basically OK" to "reinforcement theory explains the data collected in operant conditioning experiments just as Ptolemy's theory explained the data on planetary movements. But PCT explains the operant conditioning data better just as Copernicus' theory explained the planetary data better".

Is that more like what you are controlling for?

If so, being shown that reinforcement theory doesn't explain the data collected in operant conditioning experiments would certainly be a disturbance, wouldn't it?

Best Rick

Date: Fri, 9 Jun 1995 13:39:27 -0500
Subject: Shootout

[From Bruce Abbott (950609.1335 EST)]

>Rick Marken (950608.2130) --

>>Bruce Abbott (950608.1930 EST)

>> I am claiming that current reinforcement [theory] provides a consistent and generally compelling framework for understanding behavior (which is why it has so many adherents) and is capable of handling the sort of data Bill P. asserts it cannot (the ratio data).

> Well, I'd say it's time for you to show how reinforcement theory handles the ratio data, pardner. I'm talking about a workin' model, friend; no curve fitten'. So it won't do ya' no good ta draw =|;-)

Black Bart downs the shot of whiskey in one gulp, calmly places the empty shot glass on the counter, and turns in the direction of the voice. There, just in front of the swinging doors that mark the entrance to the Longbranch saloon, stands yet another pimply-faced kid, his right hand poised over the cheap six-gun protruding from a shiny new brown leather holster. Bart has seen it all before-- when you're the top gunfighter every kid who wants to make a name for himself comes lookin' for you. Bart slowly reaches down and releases the strap on his own holster.

"Kid," he says, "do ya really want to do this? You've got your whole life ahead of you. Why throw it all away?"

The kid flexes his fingers over the handle of his six-gun. "Yeah, I really want to do this. I hear you're the best, and I aim to prove I'm better."

Bart shakes his head slowly back and forth. "That's what that kid in Tombstone said. He said I wouldn't even get off a shot, but I drilled him between the eyes quicker n' he could say 'e. coli.' " Bart pats the well-oiled Colt Special at his side. "Me an' the "reinforcer" here never miss. I suggest you think it over--while you still can."

Rick:

>>> I think Bruce is claiming that PCT is better (more accurate, simpler, etc) than reinforcement theory as a model of behavior change. In other words, reinforcement theory (like Ptolemaic theory) is basically OK but PCT (like Copernican theory) is much better.

> You disagreed with the "OK" part. So I will change " reinforcement... is basically OK" to "reinforcement theory explains the data collected in operant conditioning experiments just as Ptolemy's theory explained the data on planetary movements. But PCT explains the operant conditioning data better just as Copernicus' theory explained the planetary data better".

> Is that more like what you are controlling for?

> If so, being shown that reinforcement theory doesn't explain the data collected in operant conditioning experiments would certainly be a disturbance, wouldn't it?

No, it wouldn't. If I thought reinforcement theory was that good I would never have been led to look elsewhere for a better explanation. It's not that PCT is "more accurate" or "simpler" than reinforcement theory, it's that PCT provides a more fundamental kind of explanation. The comparison is more like Ptolemy versus Newton than Ptolemy versus Copernicus.

My interest in the ratio data springs from a different and more limited concern. I believe that it is possible to come up with an explanation for the ratio data that would be consistent with reinforcement theory as it is currently conceived.

Beliefs range from those about which you care little to those you would strongly defend, perhaps unto death. If I were to discover that reinforcement theory (as _I_ conceive it) could not handle the ratio data, I think my response would be "Hmmm. Well I'll be. I wonder if anyone in EAB realizes this."

What does seem apparent is that people in EAB who have looked at these data don't seem to find them particularly surprising, so I'm guessing that there is some fairly straight-forward way reinforcement theory can handle the data. Or perhaps no one has really examined the question closely enough to discover that it can't. Perhaps it _seems_ that it can when thought about on a purely verbal level. What I'm concerned about is Bill P.'s suspicion that it is understood that reinforcement theory can't handle these data but that this fact is being hidden in order to preserve the theory. I think this is highly unlikely. More importantly, in terms of overall strategy I think it destroys one's credibility to assert that a competing theory can't handle certain data, only to be shown by the "other side" that it can. This suggests that the wise course of action is to apply reinforcement theory to the situation and see what it can do.

So, I'll give it a shot, but don't expect quick results, I'm very busy with other things at the moment.

> "Now." [said Deep Thought.] "Ask what else of me you will that I may function. Speak." They shrugged at each other. Fook composed himself. "O Deep Thought computer," he said, "the task we have designed you to perform is this. We want you to tell us . . ." he paused, "the Answer!" "The Answer?" said Deep Thought. "The Answer to what?" "Life!" urged Fook. "The Universe!" said Lunkwill. "Everything!" they said in chorus. Deep Thought paused for a moment's reflection. "Tricky," he said finally. "But can you do it?" Again, a significant pause. "Yes," said Deep Thought. . . . . "But the program will take me a little while to run." Fook glanced impatiently at his watch. "How long?" he said. "Seven and a half million years," said Deep Thought.

[From Douglas Adams, _The hitchiker's guide to the galaxy_.]

Regards, Bruce

Date: Fri, 9 Jun 1995 14:33:42 -0600
Subject: Re: reinforcement theory;

[From Bill Powers (950609.0910 MDT)]

Bruce Abbott (950608.1930 EST) --

> I am claiming that current reinforcement [theory] provides a consistent and generally compelling framework for understanding behavior (which is why it has so many adherents) and is capable of handling the sort of data Bill P. asserts it cannot (the ratio data).

Compelling, yes, consistent within a narrow set of observations, perhaps, but consistent with all we know, no.

Have you ever seen one of those lawn ornaments with a windmill on it, and a little jointed man with his arms attached to a crank that turns with the spinning of the windmill? It takes very little effort, and initially, none, to see the little man as turning the crank to make the windmill go. Based on appearances only, we could make a "compelling" case for saying that the little man is turning the windmill, which is causing the wind to blow. The little man turns the crank faster to make more wind blow, turns it slower to make less wind, and stops to make the wind stop.

Since that explanation seems to fit the facts as well as the other one, why don't we believe it? Not because of what we observe the little man doing, but because the explanation doesn't fit ALL the facts we know about, or can find if we look. But first we have to be willing to look: we must admit that seeing the little man as the cause is only one possible interpretation and deliberately consider alternative interpretations.

> I need to understand how the PCT model of the universe accounts for the apparent motions of the planets, even though those motions are only a side-effect of not being at the center of the universe.

Interesting point. The PCT model says that the only universe we know is the one we perceive from our own point of view, and that this universe, while probably related to a real universe, is not necessarily a literal representation of the real one.

We observe planets apparently moving among the stars, with two of them never getting very far from the sun, and three others following paths that actually loop back on themselves every year, by different amounts. If we simply believed our observations, we would have to come up with a very complex explanation for these different irregular movements. However, when we make a model of the physical solar system, looking for underlying mechanisms that might produce these appearances, we come up with a completely different picture that doesn't look at all like what we see with our eyes. A few simple underlying principles such as the inverse-square law of gravitation and conservation of angular momentum lead to a model with planets in elliptical orbits, with us being on one spinning planet and observing the others from a moving platform.

If we believe that the model gives a more correct picture than our observations, we see that while the observations are consistent with the model, certain of them are illusions -- the apparent yearly reversal of movement of the outer planets, for example, is not real. The apparent change of shape of the other planets, as well as our own Moon, is not real. So we begin to understand that the world we perceive and control may be quite different from the world that exists.

Glad you brought it up, although I may have taken the argument in a different direction from what you had in mind.

Rick Marken (950608.2130) --

> Well, I'd say it's time for you to show how reinforcement theory handles the ratio data, pardner. I'm talking about a workin' model, friend; no curve fitten'. So it won't do ya' no good ta draw =|;-)

Let's make it simpler. Here are two experimental points:

Ratio	reinforcement rate	behavior rate
1	210	210
40	90	3000

These are estimates from a printed graph, representing the average for four rats over an entire session. Same old graph, same old source.

If an increase in reinforcement causes an increase in behavior, how come a decrease in reinforcement rate of 120 causes an increase in behavior rate of 2790 (both per session)?

From the same graph, we have

Ratio	reinforcement rate	behavior rate
40	90	3000
160	<10	1200

Now, for a decrease in reinforcement rate of 80+, we get a decrease in reinforcement rate of 1800. This is the "right" relationship. Again, these are estimates from a printed graph, but the errors are not likely to be more than 10% or so.

Notice that the "right" relationship holds only for the lower 90/210 or 44% of the range of reinforcement rates, and that the relationship between the extremes is

Ratio	reinforcement rate	behavior rate
1	210	210
160	<10	1200

It certainly looks as if the most general rule is that the less reinforcement there is, the more behavior there is. Only if you keep the received reinforcement rate in the lower range will the opposite or "right" relationship be seen in terms of incremental changes.

Best to all, Bill P.

Date: Fri, 9 Jun 1995 15:21:41 -0700
Subject: Black Bart and The Kid

[From Rick Marken (950609.1530)]

Bruce Abbott (950609.1335 EST) --

> "That's what that kid in Tombstone said. He said I wouldn't even get off a shot, but I drilled him between the eyes quicker n' he could say 'e. coli.' "

I heard about that kid. I think you'd best speak with the Tombstone coroner, a fella named Bill Powers, about what actually happened in that standoff. I seem to recall hearing about a smart, good-lookin' cow puncher who hog-tied a big, drunken varmint who kept wandering randomly around town screaming "reinforcement", "reinforcement". Funny how these stories get distorted;-)

> I believe that it is possible to come up with an explanation for the ratio data that would be consistent with reinforcement theory as it is currently conceived.

And I believe for every drop of rain that falls a flower grows.

It seems to me that it would be rather crucial to TEST that belief, don't you?

> the wise course of action is to apply reinforcement theory to the situation [ration schedule experiments] and see what it can do.

> So, I'll give it a shot, but don't expect quick results, I'm very busy with other things at the moment.

You sure your not that Canadian fella? He always said EXACTLY the same thing about information theory -- and we have still seen no results. Seems like I'd want to know _pretty quick_ whether or not a theory that I had always assumed was correct could account for the data on which it was based!

But maybe it's not really that important, eh? After all, it's just a theory, right;-)

Best Rick

Date: Fri, 9 Jun 1995 17:43:23 -0600
Subject: Re: Judge Ito; demonstrating reinforcement

[From Bill Powers (950609.1600 MDT)]

Bruce Abbott (950609.1145 EST) --
Rick Marken (950608.2130) --

The Judge Ito metaphor [see Bruce Abbott (950608.1750) in SHOCK_EX.PER] is appropriate to the argument between you and Rick. In this system of determining truth, each side takes a position and then looks selectively for all possible evidence to support it, while also seeking to hide or suppress all evidence against it.

This is not, of course, how science is supposed to work, although it often does work this way.

Probably the least important "scientific" argument of all goes like this:

"Thousands of brilliant people have believed in my position, including Nobel Prize winners and other people of impeccable scientific credentials. To say that this position is wrong is to say that all these people have been wrong, all of this time. That is so unlikely that we have to assume it can't be true. Do you really think that you can find some flaw that these skilled scientists have not already considered and dealt with? Are you saying you are right, and all these thousands and thousands of scientists are wrong?"

The problem with this argument is twofold.

First, thousands and thousands of scientists are even more likely to be wrong than a single scientist, because most of them are simply following the leaders, adopting their interpretations cookbook style and repeating the catchy phrases they make up. If they copy the experiments, they also copy the interpretations. They are also likely to be anywhere from a month to a decade out of date with respect to what the leaders are now saying. Furthermore, only a tiny handful of scientists will actually have done the original experiments and the original interpretation of the data; all that the rest of them know about is whatever conclusions were published. So if the original handful was in error in some way, the main result of spreading their ideas around is to amplify the error. The likelihood of discovering the error decreases as the number of scientists who believe there can't be any error increases.

Second, this argument says absolutely nothing about the substance of whatever position is being taken: the validity of the data, the validity of the interpretation, or the validity of the application to any real case.

This is, of course, the principle of the Expert Witness. The more expert witnesses you can call and the more impressive their credentials, the more the jury is presumably impressed with the correctness of your position. Unfortunately the other side counters with an equal or greater number of expert witnesses with equal or better credentials and in support of exactly the opposite position. While this should tell us something about expert witnesses, apparently it does not.

Scientific beliefs are determined by a vote. Scientific truths are not.

Rick tells me, Bruce, although the post hasn't reached here, that you have referred back to the E. coli programs to show a successful application of the principle of reinforcement. I rather suspected you would, and hoped you wouldn't.

The problem with that demonstration is that the basic model had already been developed, so most of the explanation of the observed behavior had been worked out. All that was left was for you to fill in a mechanism whereby discriminative stimuli and reinforcements would lead to the right answer, which was already known: the organism should tumble sooner when going down the gradient, and later when going up it. The role of the period of straight-line swimming was also already built into the model.

Since that was already known, you could call any variable you liked a discriminative stimulus, and say that any change of conditions you liked was a reinforcer. The only constraint was that the outcome must be a shortening of the delay before a tumble when going the wrong way, and so forth. With this amount of freedom, almost any model at all could be made to fit the behavior. Since you knew that when swimming the wrong way the probability of a tumble per unit time must increase, you knew that whatever reinforcer there was, it should increase that probability. Then all you had to do was see what conditions held among the other variables when swimming went the wrong way, and assign them "reinforcing" or "punishing" properties as needed. Those properties included the ability to change a probability of a tumble, but how that was accomplished was not modeled. The effect was just calculated.

In our model of e. coli, there was no assumed component that required us to postulate that control exists. There were no nonphysical mechanisms proposed. The OUTCOME of the model's organization was that it produced control, but the capacity to control was not among the premises. This is the difference between your approach and ours: yours required building in the very phenomenon whose existence you are trying to demonstrate.

I had a great deal of trouble understanding your e. coli model. I now realize that it was because I was looking for the origin of the reinforcing effect or the discriminative effect, and couldn't find it -- yet the model seemed to work. So subtly was the question begged that the point where it was begged slid right past me unnoticed.

If we forget about labels such as reinforcer and discriminative stimulus, what we are left with in your model is a perceptual system that detects a logical condition involving two variables, and a mechanism involving continuous calculation of probability density as the means for altering the delay between tumbles, along with a link that can change the probability density. Of the four possible logical conditions, two have an effect in the wrong direction, but are outweighed by the other two. If we eliminate the inappropriate connections we are reduced to a simple on-off perceptual system, and can combine the two separate mechanisms for altering the delay between tumbles into one simpler one. After substituting for your magical connection a more physiological mechanism for changing the delay between tumbles, we are left with a simple control system of the kind that Rick and I had proposed.

My conclusion is that your e. coli model did not demonstrate that reinforcement is a real phenomenon; it simply assumed that it was.

I hope that in the case of the simple data tables I posted earlier, your argument in favor of the reinforcement interpretation will not assume that this interpretation is correct as a premise of the argument.

Best, Bill P.

Date: Sat, 10 Jun 1995 04:59:29 -0600
Subject: The E. coli reinforcement models

[From Bill Powers (950610.0300 MDT)]

Bruce Abbott (9506xx) --

If you remember that your Ecoli programs were a successful model of reinforcement theory, then either I didn't communicate properly or you are misremembering. I think I can do better now.

Here is a block diagram of Ecoli3:

```
                            dN* = 0 (reference signal)
                             |              S+
               S             - --------------------------
     ---------------------->| |      S-                  |
    |                        - ----------------          |
    |                                          |         |
    |                                          |         |
S = | dN/dt                               [inc PTS-] [dec PTS+]
    |                                          |         |
[Input F = d/dt]                          PTS- ->    PTS+ ->
    |                                       delay      delay
    |                                          \         /
    | N                                         \      /
    |                                            TUMBLE
    |                                               |
    |                                           direction
    |                                               |
     --------<------[environment]--------<----------
```

PTS+ == probability of tumble given S > reference (S+),
PTS- == probability of tumble given S < reference (S-).

If the nutrient rate of change dN/dt is negative, then PTS-, each time it is used, is made larger and PTS- is also used to determine the delay before the next tumble. If dN/dt is positive, then PTS+, each time it is used, is made smaller and is also used to determine the delay. So PTS- always goes toward max and PTS+ always goes toward min, and PTS- is always chosen when going down the gradient and PTS+ is always chosen when going up the gradient. The delay if going the wrong way is always shorter, and if going the right way is always longer. In other words, if there is any "learning" in this system, what is to be learned is already built in, and the same thing will be "learned" regardless of circumstances. The changes in PTS+ and PTS-, and their limits, are predetermined, not learned. All you did was start PTS+ and PTS- at equal levels, and then made sure they could only change, and would in fact change, toward the required levels.

In going to Ecoli4, you may have realized this deficiency, and also that there was only a discriminative stimulus and no reinforcer in this model. The block diagram of Ecoli4 now includes a "reward-or-punish" routine, which uses the second derivative of N as the reinforcer R.

Here is a block diagram of Ecoli4:

```
                            dN* = 0
                             |              S+
               S             - ----------->--------------
     ---------------------->| |       S-                 |
    |                        - ----------->----          |
    |                                          |         |
    |                                    R-    *         *   R-
    |-->--[d/dt]-> R = (dS/dt > 0)-->     --- /  switch    \ ---
    |                       operates    |      | R+   R+ |     |
S = | dN/dt                  switch    dec    inc       dec   inc
    |                                  PTS-   PTS-      PTS-  PTS-
    |                                     \    |         |   /
    |                                       \  |         | /
[Input F = d/dt]                 Output    [PTS- ->   [PTS+ ->
    |                            Function    delay]     delay]
    |                                          \        /
    | N                                          \      /
    ^                                             TUMBLE
    |                                               |
    |                                           direction
    |                                               |
     ---------------[environment]-------------------
```

S = discriminative stimulus
R = reinforcement: + = rewarding, - = punishing
N = nutrient

When both poles of the switch are vertical, the overall function is exactly the same as in Ecoli3. In other words, it is predetermined that delays will be short when going the wrong way, and long when going the right way. Reaching the maximum effect is slowed, but reaching the correct effect is inevitable and built in beforehand.

When the change in dN/dt across a tumble is negative, however, the switch will be thrown so that PTS- will be decremented just before being used to determine the delay, and PTS+ will be incremented just before being used. These changes are in the wrong direction for progressing up a gradient. Therefore the added features of Ecoli4 reduce the speed with which it will approach the fastest progress up the gradient. The reinforcement effect makes "learning" slower than it would be without the reinforcement effect.

The only reason that adding the reinforcement path does not destroy control completely is that the environmental geometry is centered on a point-source of nutrient. If the gradient did not converge toward a point (if it behaved as for a line source or a very distant point source), the probability of the second derivative being positive would be 50% regardless of the value of the first derivative. Then the switch would spend as much time in the wrong position as the right position, and PTS+ and PTS- would wander at random.

Ecoli3, on the other hand, would progress up a nonconvergent or even a divergent gradient as usual.

What makes your model work is the fact that it varies the delay appropriately to control the rate of change of nutrient, keeping it positive. The reason it does that is that it was modeled directly on the successful control-system model. The logic of Ecoli3 is identical to the logical structure of the control system model, although the mechanisms for converting errors into appropriate delays are unnecessarily complex. And in that successful model, there is nothing that corresponds to the idea of reinforcement.

Your judgment that the reinforcement model "works" was based only on the fact that the resulting behavior was correct: E. coli did approach the target. But I have spent a lot of time in careful analysis of the logic of your model, trying to understand why it does work, not just that it does work. And I have found that what makes your model work has nothing to do with reinforcement theory: adding explicit reinforcement in the way you did _worsens_ the performance of the model.

When you were arguing that your model did work, you did not go through the model as I have done to see whether it worked as you said it worked. You went rapidly through some verbal arguments, but the clincher for you was that the right result occurred: E. coli approached the target. I think it would be instructive to speculate about why your verbal arguments seemed sufficient, when in fact they glossed over fundamental defects in the logic. I think it would be reasonable to say that you simply couldn't believe that reinforcement theory would not work. Assuming that it had to work, you didn't see any reason to go through the details of your system and figure out what it would actually do according to its own structure, instead of according to what you expected and wanted it to do. Your reasoning, in fact, was driven by the goal, being adjusted to make the perception match the goal-perception.

Obviously, I don't consider this to be a sin. It is a very common phenomenon that goes a long way toward explaining the phenomenon of belief. Belief is not just a passive perceptual phenomenon; it's an active control process in which inputs are selected that will support the belief. When we have reason to want a conclusion to be true, we can construct logical arguments that are just complete enough to support the desired result, but not so complete as to risk disproving it. We find this phenomenon everywhere in human affairs, in science and everywhere else. The only thing that permits science to exist at all is the fact that there are some who wish for different conclusions, and select different logic to support their belief. In the ensuing conflicts and arguments, all sides are forced to look at their models in more detail to see what they would actually do rather than what they are believed to do. So all sides are eventually brought under the discipline of mathematically correct logic, whether they want to be or not.

Best, Bill P.

Date: Sat, 10 Jun 1995 09:07:09 -0700
Subject: Re: The E. coli reinforcement models

[From Rick Marken (950610.0900)]

Bill Powers (950610.0300 MDT) to Bruce Abbott (9506xx) --

> If you remember that your Ecoli programs were a successful model of reinforcement theory, then either I didn't communicate properly or you are misremembering. I think I can do better now.

You are obviously at your best VERY early in the morning, Bill. This was a very clear and helpful post -- both about the "reinforcement" model and about the nature of belief itself. In the interests of keeping the dialog on this VERY important topic as clear as possible, I will try to restrain myself until Bruce has had a chance to reply. And I am still looking forward to seeing his reinforcement model of the ratio schedule data you posted (Bill Powers (950609.0910 MDT)).

Best Rick (The Kid) Marken

Date: Sat, 10 Jun 1995 20:20:46 -0500
Subject: Interesting Point

[From Bruce Abbott (950610.2020 EST)]

> Bill Powers (950609.0910 MDT)] --
> Bruce Abbott (950608.1930 EST)

>> I am claiming that current reinforcement [theory] provides a consistent and generally compelling framework for understanding behavior (which is why it has so many adherents) and is capable of handling the sort of data Bill P. asserts it cannot (the ratio data).

> Compelling, yes, consistent within a narrow set of observations, perhaps, but consistent with all we know, no.

Note that in the above quote I have made the narrowest of claims for reinforcement theory: that it can handle the ratio data. I don't see where I claimed that it was "consistent with all we know." Please point out that part so I can correct it.

>> I need to understand how the PCT model of the universe accounts for the apparent motions of the planets, even though those motions are only a side-effect of not being at the center of the universe.

> Interesting point. The PCT model says that the only universe we know is the one we perceive from our own point of view, and that this universe, while probably related to a real universe, is not necessarily a literal representation of the real one.

> . . .

> If we believe that the model gives a more correct picture than our observations, we see that while the observations are consistent with the model, certain of them are illusions -- the apparent yearly reversal of movement of the outer planets, for example, is not real. The apparent change of shape of the other planets, as well as our own Moon, is not real. So we begin to understand that the world we perceive and control may be quite different from the world that exists.

Amen. Dag, here's another one for your wall (seriously!).

> Glad you brought it up, although I may have taken the argument in a different direction from what you had in mind.

Well, it's really a different argument entirely. I don't mind your using the example for a different purpose so long as my argument does not get lost in the process. I didn't see any response to it: did I miss the reply? Does anyone remember what the argument was?

Regards, Bruce

Date: Sat, 10 Jun 1995 21:06:44 -0500
Subject: Legal Maneuvers

[From Bruce Abbott (950610.2105)]

>Bill Powers (950609.1600 MDT)]

>Bruce Abbott (950609.1145 EST) --
>Rick Marken (950608.2130) --

> The Judge Ito metaphor is appropriate to the argument between you and Rick. In this system of determining truth, each side takes a position and then looks selectively for all possible evidence to support it, while also seeking to hide or suppress all evidence against it.

I see that the significance of my little parody has not been lost. I have been feeling for some time now that I have been trying to argue science with a lawyer, so I thought we'd just move into the courtroom. At least that style of argumentation is appropriate there.

Although you are certain to disagree, I have throughout these exchanges avoided (at least to the best of my ability) such behavior. If you think I have been doing otherwise, please give examples where I have slipped up, so that I can spot when I'm making these mistakes in the future and correct them.

> This is not, of course, how science is supposed to work, although it often does work this way.

Unfortunately true.

> Probably the least important "scientific" argument of all goes like this:

> "Thousands of brilliant people have believed in my position, including Nobel Prize winners and other people of impeccable scientific credentials. To say that this position is wrong is to say that all these people have been wrong, all of this time. That is so unlikely that we have to assume it can't be true. Do you really think that you can find some flaw that these skilled scientists have not already considered and dealt with? Are you saying you are right, and all these thousands and thousands of scientists are wrong?"

This is not an argument _I_ ever offered (I'm well aware of its falsehood) and I don't recall Rick making it, so I'm puzzled: are you just offering an example of invalid argument too often seen in "scientific" debate, or were you suggesting that one of us has actually proposed such an argument? From this end it's a bit ambiguous.

Regards, Bruce

Date: Sun, 11 Jun 1995 11:29:39 -0500
Subject: Shootout, Continued

[From Bruce Abbott (950611.1125 EST)]

As we return from the commercial break, Black Bart is standing, his back to the bar, facing the saloon doors through which his young challenger had entered a few moments before. Bart's eyes are calmly fixed on the kid, who is posed in a gunfighter's stance, his right hand nervously twitching above the handle of a cheap six-gun. As the moments pass, the kid's hand begins to tremble. His nerve gone, the kid turns suddenly and bolts through the saloon doors, swinging them hard enough to slam them against their stops. The kid mounts his horse and rides off at a full gallop.

Bart turns back to the bar. "Bartender!" he says, "Gimme another shot of whiskey."

Scene II: Back at the Ranch

Bill McGraw and his wife are just finishing up their evening meal when they notice the sound of an approaching horse being ridden hard. The hoof-beats get louder and then stop; suddenly the front door of the ranch-house swings open and in bolts the kid, his face flush with rage and embarrassment, his eyes glistening with tears.

"Paw!" he says, "It ain't fair! He used the 'e-word' on me!"

McGraw rises from the table and reaches for the gunbelt hanging from the back of his chair. "I thought he might, but I hoped he wouldn't."

McGraw walks to the door, swings it open, then turns his face back toward the kid and his Ma. "Ya done good, kid, but this is a man's job. Black Bart is trickier 'n a riverboat gambler 'n twice as low. I'll take it from here."

Scene III: The Long Branch Saloon

Bart is sipping on his second whiskey and chatting with one of the "ladies of the evening" when a loud voice carries over the sound of the honky-tonk piano.

"Bart! I want to have a little talk with you!"

There stands Bill McGraw, now using his Stetson to beat the dust out of his clothes. "I understand you been spreadin' lies about that fight we had back in Tombstone. Either you got a bad memory or you completely misunderstood what happened. I'm here to straighten you out."

Bart looks McGraw right in the eye. "Yes, I think we do need to talk this over. Bartender! Pour my friend Bill here a drink! Lets sit down and see if we can work this out peaceable-like."

To Be Continued . . .

Date: Sun, 11 Jun 1995 12:32:01 -0500
Subject: High Gain

[From Bruce Abbott (950611.1230 EST)]

If a control system has very high gain then very small errors can produce very large outputs. I'm wondering about the implications of the following data:

Disturbance: "e. coli" (two words)

Reference: zero words (inferred from "I hoped you wouldn't [say that].")

Error: 2 - 0 = 2 words

System Output: 615 counter-words (1st post) + (1125 counter-words (2nd post) = 1740 counter-words

Apparent Gain: 1740/2 = 780

Looks like the reference is well-defended, but I'd be a little worried about overshoot, given the large transport lag in the system: this much output before there is any apparent effect on the controlled variable! (;->

Regards, Bruce

Date: Sun, 11 Jun 1995 14:32:52 -0700
Subject: Away with words

[From Rick Marken (950611.1430)]

Bill Powers (950609.1600 MDT) --

> Probably the least important "scientific" argument of all goes like this:

> "Thousands of brilliant people have believed in my position...To say that this position is wrong is to say that all these people have been wrong, all of this time. That is so unlikely that we have to assume it can't be true.

Bruce Abbott (950610.2105) --

> This is not an argument _I_ ever offered

Bruce Abbott (950609.1335 EST) --

> What does seem apparent is that people in EAB who have looked at these data don't seem to find them particularly surprising, so I'm guessing that there is some fairly straight-forward way reinforcement theory can handle the data.

Best Rick

Date: Sun, 11 Jun 1995 16:37:27 -0500
Subject: E. Coli Revisionism

[From Bruce Abbott (950611.1635 EST)]

>Bill Powers (950609.1600 MDT) --

> Rick tells me, Bruce, although the post hasn't reached here, that you have referred back to the E. coli programs to show a successful application of the principle of reinforcement. I rather suspected you would, and hoped you wouldn't.

I think you have serious misunderstandings about why I wrote those programs, and what I believe I was able, ultimately, to demonstrate. I hope I can clear these misunderstandings up.

Be forewarned that I intend to be as of critical of your reasoning as you have been of mine. I'm not doing this to be nasty, but since you have felt compelled to "call 'em as you see 'em," I have felt at liberty to do the same. Fair is fair.

I. The Challenge

The challenge was to build a model, consistent with reinforcement theory, that would behave "properly" under special conditions designed to thwart a reinforcement interpretation. Those conditions were as follows:

1. The only behavior permitted was straight-line motion or "tumbling."

2. The outcome of a tumble was to sample a new direction for straight-line motion. The new direction following a tumble was to be chosen strictly at random.

3. The only sensory inputs available were nutrient density and its derivatives.

4. The test environment would consist of a two-dimensional field and, located in that field, a nutrient source, with nutrient concentration declining with distance in every direction away from the source.

The model would be considered successful if it climbed the nutrient gradient and then tended to stay on or near the nutrient source.

II. The Purpose of the Demonstration

The purpose of the demonstration was to prove that a model based on reinforcement principles _could be constructed_ which would behave as specified (a proof of principle). It was asserted that one could not.

III. What was Not the Purpose of the Demonstration

It was NOT the purpose of the demonstration to show that

1. reinforcement-based and PCT models of e. coli behavior are equally "valid," robust, explanatory.

2. reinforcement is a real phenomenon.

To return to our favorite illustration, asserting that NO reinforcement-based model could behave properly in the test situation is equivalent to asserting that the Ptolemaic system could not properly describe the motions of Mars. Ptolemaic theory may be wrong, but is it true that it can't handle these data?

I am arguing that your assessment of the inability of Ptolemaic theory to handle the Mars data is incorrect. I am NOT arguing that Ptolemaic theory is correct! NOR am I arguing that it can handle ALL data.

Big difference, but one that seems to have been missed. You say:

> My conclusion is that your e. coli model did not demonstrate that reinforcement is a real phenomenon; it simply assumed that it was.

Yes, just as my calculations based on cycles and epicycles would not have demonstrated that cycles, epicycles, and retrograde planetary motion are real phenomena; they would only have demonstrated what I set out to demonstrate, that such calculations do provide an account of planetary motion that is consistent with their observed paths through the night sky.

As I did not set out to demonstrate that reinforcement is a "real" phenomenon, I cannot be faulted for failing to demonstrate it.

IV. Consistency of ECOLI4a with reinforcement theory.

In attempting to develop a model that would perform according to criterion in the specified, very unfavorable (with respect to the theory) test environment, I considered several approaches before arriving at one that appeared to work. This model incorporated several standard reinforcement-theory concepts, as described below:

1. Reinforcement
   This is an increment in the probability of a response when the response is immediately followed by a favorable outcome, in this case, an increase in the rate of nutrient increase.

2. Punishment
   This is an decrement in the probability of a response when the response is immediately followed by an unfavorable outcome, in this case, a decrease in the rate of nutrient increase.

3. Stimulus Control
   The response probabilities affected by reinforcement/punishment are conditional probabilities; the conditions are identified with discriminably different stimulus conditions. Here, we defined a rising nutrient gradient as S+ and a falling nutrient gradient as S-.

For these relationships to hold, the organism must be assumed to have appropriate structures that provide the necessary functions. For reinforcement and punishment to work, there must a sensory structure that detects the rate of change in nutrient concentration, a structure to store the rate immediately prior to a tumble, and a structure to compare this rate to the rate immediately after a tumble. It is not difficult to imagine a set of molecular components that might provide these functions, but these would involve pure speculation on my part, so I included in the model only what they do, not how they do them. For discrimination to take place, there would also need to be a mechanism that could selectively associate the stored state of nutrient change prior to a tumble with the appropriate structural representation of tumble probability (perhaps the concentration of a chemical whose effect on tumble probability is mediated by an enzyme whose concentration represents the stored value of, say S+, but again, such mechanisms are speculative; only the functions are modeled).

The model works as follows:

1. When S+ was present prior to a tumble and the tumble is reinforced, the probability of a tumble given S+ is increased.

2. When S- was present prior to a tumble and the tumble is reinforced, the probability of a tumble given S- is increased.

3. When S+ was present prior to a tumble and the tumble is punished, the probability of a tumble given S+ is decreased.

4. When S- was present prior to a tumble and the tumble is punished, the probability of a tumble given S- is decreased.

All this is consistent with common principles of reinforcement theory.

V. How the Model Performed in the Test Environment

Because the test environment contained a point source of nutrient, tumbles made while moving "upstream" (up the nutrient gradient) yielded a more favorable rate of nutrient change (reinforcement) on only 1/4 of tumbles and a less favorable rate of nutrient change (punishment) on 3/4 of tumbles. Over the long run this drove the probability of a tumble in the presence of S+ down to its minimum value. E. coli virtually stopped tumbling when moving up the nutrient gradient.

The reverse condition held when e. coli was moving down the nutrient gradient, so that the probability of a tumble in the presence of S- was driven upward to its maximum value. E. coli tumbled frequently when moving down the nutrient gradient.

The result of these two changes in tumble probability was that e. coli behaves as required by the Challenge. Given the right set of cycles and epicycles, the Ptolemaic system can indeed reproduce the apparent motions of Mars.

VI. Replies to Specific Criticisms [Bill Powers (950610.0300 MDT)]

Re: ECOLI3

This model was withdrawn after I realized that I had made some errors of application. End of discussion.

Re: ECOLI4

> When both poles of the switch are vertical, the overall function is exactly the same as in Ecoli3. In other words, it is predetermined that delays will be short when going the wrong way, and long when going the right way. Reaching the maximum effect is slowed, but reaching the correct effect is inevitable and built in beforehand.

I am having trouble following your diagram but I believe it is wrong. (I can't make out what those switches are doing, where the pivot is. In addition some of the critical elements appear to have been mislabeled.) The mechanism I have given e. coli has no "predetermined," "correct," or "inevitable" direction of change. It is only the nature of the environment that causes the two tumble probabilities to move in the "correct" (i.e., adaptive) directions.

We went over this in some detail a few months ago and the end result was that you (finally) got it right. This diagram suggests that you are confused again.

You may be speaking of the effect of reinforcement and punishment on tumble probabilities (which is the same whether S+ or S- is present). In that case your complaint is that I have built the model so that reinforcement reinforces and punishment punishes. This hardly seems a valid complaint.

The analogous complaint for the control model is that you have given it negative gain. Why should gain be negative? Why, because if it isn't, e. coli won't move up the nutrient gradient! Put another way, you assume in the control model that e. coli "wants" to see nutrient levels increasing. I assume the same thing in my code for the effects of reinforcement.

> The only reason that adding the reinforcement path does not destroy control completely is that the environmental geometry is centered on a point-source of nutrient. If the gradient did not converge toward a point (if it behaved as for a line source or a very distant point source), the probability of the second derivative being positive would be 50% regardless of the value of the first derivative. Then the switch would spend as much time in the wrong position as the right position, and PTS+ and PTS- would wander at random.

True. So what? The Challenge was not to build a model that would work in any gradient, it was to build one that would work in the gradient supplied. This model does. Whether it works in other environments is irrelevant.

> Your judgment that the reinforcement model "works" was based only on the fact that the resulting behavior was correct: E. coli did approach the target.

That was the criterion for success of the model. As I recall, it was no trivial exercise to find a model that did "work." But the model had other criteria as well: it had to be consistent with the principles of reinforcement theory, and it was.

> But I have spent a lot of time in careful analysis of the logic of your model, trying to understand why it does work, not just that it does work. And I have found that what makes your model work has nothing to do with reinforcement theory: adding explicit reinforcement in the way you did _worsens_ the performance of the model.

I believe I have shown (above) that the model is entirely consistent with reinforcement theory. It is true that on 1/4 of tumbles the appropriate conditional probabilities get moved in the "wrong" direction by the model, given the point source of nutrient. That's the breaks: the model knows nothing about the effect of its behavior except what it can learn about these from the appropriate comparisons. Sometimes what it learns on a given trial is just plain wrong when viewed against the long-term interests of the bacterium. Because many "trials" give false information, it is important that the probabilities not be adjusted too much after a given trial: this would lead to instability. By making the adjustment process move slower than the process it is adjusting, we prevent instability. I believe you've pointed to the same necessity when speaking about reorganization, and it's certainly built into Hans Bloom's adaptive controller. There are some real lessons you might have learned here about the requirements of an adaptive system, but you've been too bent on demolishing the model to notice.

> When you were arguing that your model did work, you did not go through the model as I have done to see whether it worked as you said it worked. You went rapidly through some verbal arguments, but the clincher for you was that the right result occurred: E. coli approached the target.

Now these are real fighting words, Bill. I suggest you get out those old posts of mine concerning ECOLI4a and READ THEM CAREFULLY. Talk about selective memory! Wow!

As I recall, I expended considerable effort carefully describing the mechanism of ECOLI4a. We went through at least two mis-descriptions on your part and I posted not only a clear diagram of the model's logic but an equally clear diagram as to how the specific nutrient gradient determined the outcome of the simulation. Remember those "Marken probabilities," you know, where Rick said the outcome HAD to be 50-50, so that no learning was possible, his computer program said so, never mind the diagram? I strongly encourage you to go review those exchanges and see whether your recollection of the events matches what appears there.

> I think it would be instructive to speculate about why your verbal arguments seemed sufficient, when in fact they glossed over fundamental defects in the logic. I think it would be reasonable to say that you simply couldn't believe that reinforcement theory would not work. Assuming that it had to work, you didn't see any reason to go through the details of your system and figure out what it would actually do according to its own structure, instead of according to what you expected and wanted it to do. Your reasoning, in fact, was driven by the goal, being adjusted to make the perception match the goal-perception.

Sorry, Bill. (1) My arguments are sufficient. (2) They do not gloss over any defects, fundamental or otherwise, in logic. (3) Therefore your speculations as to why I thought they were sufficient are moot. I thought they were sufficient because they are sufficient, not because I was being led by the nose by any foregone conclusions.

> Obviously, I don't consider this to be a sin. It is a very common phenomenon that goes a long way toward explaining the phenomenon of belief. Belief is not just a passive perceptual phenomenon; it's an active

control process in which inputs are selected that will support the belief. When we have reason to want a conclusion to be true, we can construct logical arguments that are just complete enough to support the desired result, but not so complete as to risk disproving it.

Bill, before you go calling the kettle black, I suggest you take a good, long look at your own performance in this little debate. You may find it to be an eye-opener. Even PCT theorists can't escape from behaving as PCT predicts. (;-
>

Regards, Bruce

Date: Mon, 12 Jun 1995 10:16:55 -0500
Subject: Re: Away with words

[From Bruce Abbott (950612.1015 EST)]

>Rick Marken (950611.1430) --

>>Bill Powers (950609.1600 MDT)

>> Probably the least important "scientific" argument of all goes like this:

>> "Thousands of brilliant people have believed in my position...To say that this position is wrong is to say that all these people have been wrong, all of this time. That is so unlikely that we have to assume it can't be true.

>>Bruce Abbott (950610.2105) --

>> This is not an argument _I_ ever offered

>>Bruce Abbott (950609.1335 EST) --

>> What does seem apparent is that people in EAB who have looked at these data don't seem to find them particularly surprising, so I'm guessing that there is some fairly straight-forward way reinforcement theory can handle the data.

You have employed selective citation to give a misleading impression. The two sentences that immediately follow the one you quoted show that I am not making the sort of argument Bill presents at all:

> What does seem apparent is that people in EAB who have looked at these data don't seem to find them particularly surprising, so I'm guessing that there is some fairly straight-forward way reinforcement theory can handle the data. Or perhaps no one has really examined the question closely enough to discover that it can't. Perhaps it _seems_ that it can when thought about on a purely verbal level.

Compare with:

> "Thousands of brilliant people have believed in my position, including Nobel Prize winners and other people of impeccable scientific credentials. To say that this position is wrong is to say that all these people have been wrong, all of this time. That is so unlikely that we have to assume it can't be true. Do you really think that you can find some flaw that these skilled scientists have not already considered and dealt with? Are you saying you are right, and all these thousands and thousands of scientists are wrong?"

To see more clearly how these arguments differ, let's boil both versions down to the bare bones:

My argument:

(a) People in EAB who have looked at the ratio data believe that they can be accounted for by reinforcement theory.

(b) They may be right, or they may be wrong.

(c) Before we state that it cannot, we ought to do some checking.

The kind of argument to which Bill P. refers, applied to this specific case:

(a) People in EAB who have looked at the ratio data believe that they can be accounted for by reinforcement theory.

(b) Because so many believe so (including noted scientists), they can't be wrong.

Perhaps I'm missing something, but these don't look like the same argument to me. Not even close.

Regards, Bruce

Date: Mon, 12 Jun 1995 12:35:26 -0500
Subject: ECOLI4a Diagram Correction

[From Bruce Abbott (950612.1230 EST)]

>Bill Powers (950610.0300 MDT) --

Having carefully studied Bill Powers' diagram for ECOLI4a, I believe that I now understand it and can identify the errors. Here is Bill's original diagram:

```
>                             dN* = 0
>                              |              S+
>            S                 - ----------->-------------
>      ---------------------->| |      S-                |
>      |                       - ----------->----         |
>      |                                        |          |
>      |                                  R-    *          *   R-
>      |-->--[d/dt]-> R = (dS/dt > 0)-->     --- /  switch    \ ---
>      |                      operates    |      | R+   R+ |     |
> S =  | dN/dt                 switch    dec    inc      dec   inc
>      |                                 PTS-   PTS-     PTS-  PTS-
>      |                                    \   |          |   /
>      |                                     \  |          |  /
>[Input F = d/dt]                 Output    [PTS- ->   [PTS+ ->
>      |                          Function    delay]     delay]
>      |                                          \        /
>      |  N                                        \      /
>      ^                                          TUMBLE
>      |                                             |
>      |                                         direction
>      |                                             |
>      ---------------[environment]-------------------
```

>S = discriminative stimulus
>R = reinforcement: + = rewarding, - = punishing
>N = nutrient

Here is the corrected version:

```
                             dN* = 0
                              |              S+
                S             - ----------->-----------------
      ----------------------->| |       S-                   |
     |                        - -------->---                 |
     |                                      |                |
     |                                      *                *
     |-->--[d/dt]-> R = (dS/dt > 0)-->    --/ --  switch   --/ --
     |                     operates   R- |      | R+   R- |      | R+
 S = | dN/dt                switch      dec     inc      dec     inc
     |                                  PTS-    PTS-     PTS+    PTS+
     |                                     \     |        |     /
     |                                      \    |        |    /
[Input F = d/dt]                 Output    [PTS- ->   [PTS+ ->
     |                           Function    delay]     delay]
     |                                          \        /
     | N                                         \      /
     ^                                            TUMBLE
     |                                              |
     |                                          direction
     |                                              |
      ---------------[environment]-------------------
```

I have restructured the switch portion of the diagram to make it clearer. The poles of the double-pole switch are depicted both facing left, as they would be if the nutrient rate had become less positive or more negative following a tumble (punishment, R-). The important change is the interchange of the R+ and R- labels for the S+ leg so that, consistent with reinforcement theory, reinforcement (R+) increments tumble probability and punishment (R-) decrements it, as was/is the case for the S- leg.

> When both poles of the switch are vertical, the overall function is exactly the same as in Ecoli3. In other words, it is predetermined that delays will be short when going the wrong way, and long when going the right way. Reaching the maximum effect is slowed, but reaching the correct effect is inevitable and built in beforehand.

In the corrected diagram, this is no longer true.

> When the change in dN/dt across a tumble is negative, however, the switch will be thrown so that PTS- will be decremented just before being used to determine the delay, and PTS+ will be incremented just before being used. These changes are in the wrong direction for progressing up a gradient. Therefore the added features of Ecoli4 reduce the speed with which it will approach the fastest progress up the gradient. The reinforcement effect makes "learning" slower than it would be without

In the corrected diagram, this is also no longer true.

> I had a great deal of trouble understanding your e. coli model. I now realize that it was because I was looking for the origin of the reinforcing effect or the discriminative effect, and couldn't find it -- yet the model seemed to work. So subtly was the question begged that the point where it was begged slid right past me unnoticed.

So much for subtle question-begging, eh? (;->

Regards, Bruce

Date: Mon, 12 Jun 1995 12:37:25 -0700
Subject: Selective Citation

[From Rick Marken (950612.1230)]

Bruce Abbott (950612.1015 EST) --

> You have employed selective citation to give a misleading impression.

Not really. I selected the citation that showed why you were having problems understanding that your E. coli model is not a reinforcement model. The relevant citation is:

> I'm guessing that there is some fairly straight-forward way reinforcement theory can handle the data.

What you call "guessing" looks a lot like "controlling" to me. While you do recognize that reinforcement theory might not be able to handle the ratio (or E. coli) data, you clearly don't want to see this happen. Indeed, you don't seem to want to see that the predictions of reinforcement theory are exactly the opposite of what is found in the E. coli and ratio studies. This, I think, is why you have avoided presenting a reinforcement model of the ratio data and continue to call your control model of E. coli behavior a reinforcement model. As Bill said, this is certainly not a sin; in fact, I think it's a good illustration of what PCT is up against, even from those who are pro-PCT.

But, of course, it is my responsibility to present the data and analyses that demonstrate that you have not yet shown that reinforcement theory can account for the E. coli phenomenon. But I will wait for (Papa) Bill to reply first because I'm so afraid of your lightning wit.

Best

Rick (Pass the Stridex) Marken

Date: Mon, 12 Jun 1995 22:29:25 -0600
Subject: E. coli model

[From Bill Powers (950612.1530 MDT)]

Bruce Abbott (950612.1230 EST) --

> Having carefully studied Bill Powers' diagram for ECOLI4a, I believe that I now understand it and can identify the errors.

My abject apologies: I can see now that I made a mistake: R+ should go with an increment of both PTSs and R- should go with a decrement of both PTSs. I should have gone back and looked at my old posts on this subject. I had remembered finding that two of the conditions were incorrect for going up the gradient, but I didn't retrace the logic this time to make sure which two they are.

But I'm not ready to let you off the hook yet. Your program's behavior does not match your description of what makes it work, and it leads to successful behavior only because of a slight imbalance between effects that work the right way and, essentially as often, the wrong way. And I think I can still show that your model would work better without the "reinforcement" you describe than with it.

Defining (NutSave > 0) = S, and

(dNut > NutSave) = (DeltaNutRate > 0) = R,

with S+ meaning S true, S- meaning S false,
R+ meaning R true, and R- meaning R false,

we can write the program in shorthand:

```
if R+ then
  if S+ then
       inc(PTS+)
  else inc(PTS-)  *
else if R- then
 if S+ then
      dec(PTS+)   *
 else dec(PTS-).
```

> The important change is the interchange of the R+ and R- labels for the S+ leg so that, consistent with reinforcement theory, reinforcement (R+) increments tumble probability and punishment (R-) decrements it, as was/is the case for the S- leg.

Note that you are describing the condition as shown in the first three lines:

```
if R+ then
  if S+ then
       inc(PTS+)
```

and also

```
if R- then
 if S- then
      dec(PTS-)
```

But neither of these is what you want. If you are currently going in the right direction (S+), you do not want to _increase_ the probability of a tumble, because that will terminate the present progress sooner. If you're going the wrong way, you don't want to _decrease_ the probability of a tumble. Your verbal description makes it sound as though increasing the probability of a tumble is good, because S+ is favorable and the previous direction was less favorable (dNut - NutSave > 0). Incrementing PTS+ is the WRONG MOVE (as is the other case, decrementing PTS-).

If we had only the starred lines, your model would converge to the optimum state very quickly (depending only on how large you made the learning constant). It could easily be made to do so in one iteration. The reason is that optimum performance depends only on getting PTS- to 1.0 and PTS+ to zero. After they reach those values, the approach to the target will be as rapid as possible. So if we just said

```
  if S- then
        inc(PTS-)
  else if S+ then
      dec(PTS+)
```

The model would converge to optimum performance without ever changing the PTS probabilities the wrong way. So the model would work much better without even considering reinforcement.

As PTS+ and PTS- approach the optimum levels, the approach of E. coli to the target becomes more and more direct, with less time spent going the wrong way. Comparing the PCT model and the reinforcement model is not best done, however, in terms of approach to the target. Once the reinforcement model has made PTS+ = 0 and PTS- = 1, the approach to the target will be as fast for either model. The comparison has to be in terms of how long it takes to get the probabilities to the required states. When you have some conditions that change the probabilities away from the required states, obviously the approach to optimum conditions will take much longer. If a model without reinforcement can produce the optimum settings faster than a model with reinforcement, then clearly reinforcement is not helping.

The problem with verbal descriptions of the E. coli effect is that the words are too loose to use in a quantitative discussion. The words "tumble" and "probability of a tumble" do not refer to the same aspect of the situation. A tumble is a random change in direction. The probability of a tumble is the chance of a tumble occurring in a given time interval. The two concepts are independent: you can have any change in direction going with the occurrence of a tumble within any time interval. Affecting the time interval does not have any effect on the subsequent direction of movement. That is, if E. coli was moving in a favorable direction before a tumble, after the tumble -- however soon or late it occurs -- the next direction of movement is completely independent of the direction before the tumble.

The verbal description says that if you're already going in the right direction, a tumble would be most likely to result in a less favorable direction of movement. So if dNut is positive, the result should be to make a tumble less likely: PTS+ should be decremented. If you're already going in the wrong direction, any tumble would likely produce an improvement: PTS- should be incremented.

However, there are two other cases: where you're going in a moderately unfavorable direction and the next tumble makes it more unfavorable, and the same for going in a moderately favorable direction. Now the change in PTS based on the change in dNut necessarily leads to the wrong change in probability.

When you introduce the _change_ in dNut across a tumble, you lose all reference to the current direction of movement. A large change in direction could result in equal values of dNut, or in either an increase or a decrease in dNut.

This means that the outcome depends on the detailed way in which favorable results are affected by different directions of movement. With some distributions, the result might be progress up the gradient; with others, regress down the gradient, the wrong way. So there is no way that a model based on discriminative stimuli and reinforcements can predict movement up the gradient. That it did in this case was a matter of luck and geometry.

Best to all, Bill P.

Date: Tue, 13 Jun 1995 11:14:34 -0600
Subject: Bruce:100; Bill and Rick: 0

[From Bill Powers (950613.0815 MDT)]

Bruce Abbott (950611.1635 EST) --

Rick and I have been wrong and you have been right.

Last night I finally did a little experiment. I set up the following program:

```
==============================
program testecol;
uses dos,crt;
{
if R+ then
  if S+ then
       inc(PTS+)
  else inc(PTS-)  *
else if R- then
 if S+ then
      dec(PTS+)   *
 else dec(PTS-).

R := n2 > n1;   S = n1 > 0;

inc(pts+) = n2 > n1 and n1 > 0
inc(pts-) = n2 > n1 and n1 < 0
dec(pts+) = n2 < n1 and n1 > 0
dec(pts-) = n2 < n1 and n1 < 0
----------------
}
```

```
var n1,n2,q1,q2,q3,q4,t,tmax: real;
    i: word;
    ch: char;
begin
 q1 := 0.0;
 q2 := 0.0;
 q3 := 0.0;
 q4 := 0.0;
 randomize;
 clrscr;
 tmax := 1e6;
 t := 0.0;
 while t < tmax do
 begin
   n1 := cos(2*pi*random);  {component of velocity to right}
   n2 := cos(2*pi*random);

  if (n1 > 0.0) and (n2 > n1) then q1 := q1 + 1.0; {inc(pts+)}
  if (n1 > 0.0) and (n2 < n1) then q2 := q2 + 1.0; {dec(pts+)}
  if (n1 < 0.0) and (n2 > n1) then q3 := q3 + 1.0; {inc(pts-)}
  if (n1 < 0.0) and (n2 < n1) then q4 := q4 + 1.0; {dec(pts-)}
  t := t + 1.0;
 end;

 writeln('inc(pts+) = ',q1*100.0/tmax:6:1,
         '%  dec(pts+) = ',q2*100.0/tmax:6:1,'%', chr(13),chr(10),
         'inc(pts-) = ',q3*100.0/tmax:6:1,
         '%  dec(pts-) = ',q4*100/tmax:6:1,'%');
 ch := readkey;
end.
==============================
```

The results of this "Monte Carlo" test with one million trials were

inc(pts+): 12.5% of the trials
dec(pts+): 37.5%
inc(pts-): 37.5%
dec(pts-): 12.5%

This is exactly what you said would happen, and what I have been trying to deny through verbal reasoning for months. You are perfectly right in saying that everything I have accused you of, I have been doing myself.

Well, I knew SOMEBODY was doing it.

The condition "N1>0" is equivalent to S+; "N2>N1" is equivalent to R+. It doesn't matter what N stands for. If any series of numbers is generated randomly within a fixed zero-centered range, the two most common conditions will be

```
 (n1 > 0.0) AND (n2 < n1)  and
 (n1 < 0.0) AND (n2 > n1)
```

One or the other of these two conditions will occur 75% of the time; one of the other two, the remaining 25% of the time. I don't know how you arrived at this result, but you were right. It doesn't matter whether you compute the velocity to the right, as above, or simply

```
N1 = random - 0.5
N2 = random - 0.5
```

(Where "random" returns a real number between 0.0 and 1.0).

The condition that occurs 75% of the time is the "right" condition, the one leading either to a decrement in probability of a tumble given S+, or an increment in probability of a tumble given S-. When we subtract the two "wrong" cases, we find that on the average, the probabilities are adjusted the right way at half the rate they would be adjusted without the added pair of conditions N2 > N1 or N2 < N1. But they are adjusted the right way and by a large margin.

II. The Purpose of the Demonstration

> The purpose of the demonstration was to prove that a model based on reinforcement principles _could be constructed_ which would behave as specified (a proof of principle). It was asserted that one could not.

That assertion was wrong. Rick and I have been assuming without proof that there is no systematic way to predict the outcome of random tumbles using the history of results before and after a tumble. In the E. coli case or any similar case this is not true.

Note that if we apply the above analysis to the PCT model, we will get exactly the same results. Since the tumbles are generated at random, it will still be true that the logical functions [(N1 > 0) and (N2 > N1)] and the other three will occur with the same probabilities, 12.5% and 37.5% as in the table above. This distribution is a property of randomly-generated numbers within a fixed range, passed through the appropriate logical filters. This has nothing to do with the operation of the PCT system, which does not make use of this fact.

> To return to our favorite illustration, asserting that NO reinforcement-based model could behave properly in the test situation is equivalent to asserting that the Ptolemaic system could not properly describe the motions of Mars. Ptolemaic theory may be wrong, but is it true that it can't handle these data?

You are right. Reinforcement theory does fit the data. It does so not because it is necessarily based on the mechanism that is actually producing the data, but because it expresses a natural law concerning random numbers, a law which the framers of reinforcement theory noticed empirically. The parallel to epicycles is quite exact, because what epicycles do is express the true idea that any waveform can be represented as the sum of a series of sine-waves with suitable phases and amplitudes. Using such a description, one can fit a curve to any series of planetary positions, regardless of the mechanism that is actually governing those positions. And while I haven't worked this out at all, it may be that reinforcement theory, using a change in a variable to define reinforcement and the prior state of that variable (or one related to it) to define a discriminative stimulus, can be fit to any purposive behavior regardless of the mechanism actually responsible for the behavior.

What's more important, I think, is that a physical system could be constructed with a logical perceptual function that worked according to the above program steps, and that it could become part of the mechanism of a control system. In other words, if a physical system is in fact organized to take advantage of the natural law concerning randomly generated numbers in a bounded range, the reinforcement model could be the correct model. If the stars and planets were in fact attached to rotating crystalline spheres, the epicycle model would be the correct model.

It's just as important, however, to realize (as you do) that the reinforcement model is not necessarily the correct model of a system even if it correctly describes the observable relationships. The PCT model of E. coli does not make use of the special property of bounded random numbers, yet it produces the same overall result. By the same token, the fact that the PCT model produces the right behavior does not automatically make it the right model. To select the right model, or at least the more correct one, we must turn to auxiliary evidence that can help us choose.

I am getting a feeling that you've been ahead of me for some time.

> For these relationships to hold, the organism must be assumed to have appropriate structures that provide the necessary functions. For reinforcement and punishment to work, there must a sensory structure that detects the rate of change in nutrient concentration, a structure to store the rate immediately prior to a tumble, and a structure to compare this rate to the rate immediately after a tumble. It is not difficult to imagine a set of molecular components that might provide these functions,

but these would involve pure speculation on my part, so I included in the model only what they do, not how they do them.

Yes, I can now see that such a perceptual function is possible. What I did not see before, or believe, was that there would be anything for it to perceive. I thought the distribution of percentages above would be equal.

Now let's talk about the output mechanism.

> For discrimination to take place, there would also need to be a mechanism that could selectively associate the stored state of nutrient change prior to a tumble with the appropriate structural representation of tumble probability (perhaps the concentration of a chemical whose effect on tumble probability is mediated by an enzyme whose concentration represents the stored value of, say S+, but again, such mechanisms are speculative; only the functions are modeled).

When probabilities are reduced to physical mechanisms for generating them they usually turn out to be something pretty simple. A device for converting the probable mean value of a signal to a signal representing that number could consist of a resistor and a capacitor.

In the case of your model, the probability generator is a program statement that is executed over and over until a tumble occurs; typically

```
if (Random < pTumbleGivenSplus) then DoTumble
```

The adjustment of probabilities is not done until a tumble finally occurs. So this program step does nothing but create a delay: the lower the probability of a tumble, the more iterations are likely to occur before the tumble occurs. Since the modeled E. coli continues to move on every iteration, the more distance is likely to be covered.

When the tumble occurs and the probabilities are to be adjusted, we have a (typical) program step

```
pTumbleGivenSplus := pTumbleGivenSplus - LearnRate;
```

The desired effect could just as easily be obtained by writing

```
DelayGivenSplus = DelayGivenSplus + LearnRate
```

because decreasing the probability is the same as increasing the delay. The mechanism for creating the delay is unimportant.

Thus using the concept of "changing a probability" is merely a way of altering the delay before the next tumble. There are many physical circuits that can create a delay that can be varied by varying a signal entering the delay generator. The circuit least likely to be found in a real system is one which literally calculates a random number and compares it with a fixed number, as in the program step

```
if (Random < pTumbleGivenSplus) then DoTumble.
```

In my model for operant conditioning, I convert an error signal to a frequency of bar-pressing by letting the error signal determine the rate at which a timer (an integrator) counts upward. When the integrator output reaches a fixed trigger level, an output event is generated and the timer is reset. This is a simple circuit that is easy to implement in neurons or biochemistry: a "variable-frequency relaxation oscillator."

In models of operant conditioning you and others have proposed, the variation in output event rate is accomplished just as DoTumble is calculated: by comparing a random number generator's output with a fixed value representing a probability. If the probability is increased, the next event is generated sooner: the event rate is increased. The net result is the same as increasing the error signal in my model: the threshold for the event generation is reached sooner.

So the concept of "the probability of a response" is converted simply to "the frequency of response generation," and the mechanism is converted from the literal generation of random numbers to the operation of a simple relaxation oscillator.

If the real system operates by using a simple relaxation oscillator, characterizing it as varying a "probability" is unnecessary and inappropriate. Applying the concept of probability relies on an analogy rather than a description of the actual system. If we have a choice between a literal probability calculation and a simple relaxation oscillator as elements of the model, but no direct evidence as to the actual mechanism, we would choose the simpler model rather than introducing complexity for its own sake. Or at least, I would.

> The Challenge was not to build a model that would work in any gradient, it was to build one that would work in the gradient supplied. This model does. Whether it works in other environments is irrelevant.

Actually, the relationships in the "testecol" program will appear in any gradient. We're talking about logical relationships, so nonlinearities that don't change the order of the random numbers make no difference.

> When you were arguing that your model did work, you did not go through the model as I have done to see whether it worked as you said it worked. You went rapidly through some verbal arguments, but the clincher for you was that the right result occurred: E. coli approached the target.

> Now these are real fighting words, Bill. I suggest you get out those old posts of mine concerning ECOLI4a and READ THEM CAREFULLY. Talk about selective memory! Wow!

Fighting words, indeed -- Champ.

> As I recall, I expended considerable effort carefully describing the mechanism of ECOLI4a. We went through at least two misdescriptions on your part and I posted not only a clear diagram of the model's logic but an equally clear diagram as to how the specific nutrient gradient determined the outcome of the simulation. Remember those "Marken probabilities," you know, where Rick said the outcome HAD to be 50-50, so that no learning was possible, his computer program said so, never mind the diagram? I strongly encourage you to go review those exchanges and see whether your recollection of the events matches what appears there.

At least you can give me a crumb of credit for continuing to worry over the problem and finally coming up with what was for me the missing fact. If you had known what the key problem was -- the actual distribution of probabilities -- you would no doubt have come up with a rigorous proof that they were as you claimed. I can claim distraction by terms like "reinforcement" and "discriminative stimulus," which sound very complex until you realize they can be reduced to (N2 > N1) and (N1 > 0), and nonphysical ideas like "probability of a response", which reduces to the rate of response generation or its reciprocal, delay to the next response.

> Sorry, Bill. (1) My arguments are sufficient. (2) They do not gloss over any defects, fundamental or otherwise, in logic. (3) Therefore your speculations as to why I thought they were sufficient are moot. I thought they were sufficient because they are sufficient, not because I was being led by the nose by any foregone conclusions.

Yes, yes, yes. Can we talk about something else pretty soon, I hope? I have been a living example of goal-directed reasoning, and what's worse have projected my own fault onto you.

> Bill, before you go calling the kettle black, I suggest you take a good, long look at your own performance in this little debate. You may find it to be an eye-opener. Even PCT theorists can't escape from behaving as PCT predicts. (;->

Come on, somebody else make a comment that is really truly wrong. I need someone to beat up on.

Bruce, I thank you for your steadfastness and keeping your temper during what must have been an extraordinarily frustrating debate. We will no doubt have more disagreements in the future, but I won't soon forget the lessons of this one.

Respectfully, Bill P.

Date: Tue, 13 Jun 1995 20:31:48 -0700
Subject: Reinforcement and Control

[From Rick Marken (950613.2030)]

Well, now that Bill Powers (950613.0815 MDT) has sold me down the river ;-), let me try to recap quickly what I think the E. coli debate is about and what has and has not been demonstrated.

I created the E. coli demo 10 years ago as a challenge to reinforcement theory. A subject was to move a dot to any one of three target positions on the computer screen. The dot moves in a straight line at a constant rate until the space bar is pressed. After a press the dot again moves in a straight line but in a new, randomly selected direction. Subjects learn quickly to press appropriately to get the dot to the desired target and keep it there.

The subject's goal directed behavior in this task seems to be inconsistent with my qualitative understanding of reinforcement theory which says that responses (bar presses) are selected by their consequences (direction of movement after a press). Random consequences should select random responses.

It seemed to me that a reinforcement theorist would be surprised by the results; reinforcement theory should predict that the dot would move in a random pattern on the screen; in fact, the dot moves to (and remains near) one position on the screen.

When I submitted a description of this experiment for publication, not one reinforcement theorist who reviewed the paper was surprised by the results; and each one had a different explanation of why the results were expected. All of these explanations, it turned out, were wrong.

Finally, at the end of last year, Bruce Abbott found a way to implement a reinforcement model of the behavior in the E. coli experiment. He did it by defining the consequence of a response as _change_ in direction after a press (Bruce did it in terms of gradient, rather than direction, but its the same thing in practice). It turns out that if the dot is moving toward the target before a press, the chances are that it will be moving less towards it or away from it after a press; if the dot is moving away from the target before a press, the chances are that it will be moving less away or towards it after a press.

In other words, a press when the dot is moving toward the target is most likely to make things worse (a change for the worse or "punishment") and a and a press when the dot is moving away from the target is likely to make things better (a change for the better or "reinforcement")

This is the fact that Bill Powers (950613.0815 MDT) just rediscovered (we had already determined that this was the case at the end of last year). It means that you can use the change in direction (or gradient) after a response to appropriately increment (inc) or decrement (dec) the probability of a response when moving toward (pts+) or away from (pts-) the target. This is because, as Bill notes:

> inc(pts+): 12.5% of the trials
> dec(pts+): 37.5%
> inc(pts-): 37.5%
> dec(pts-): 12.5%

So a reinforcement model of the subject's behavior in the E. coli experiment does produce the appearance of goal directed behavior when reinforcement is defined as change in dot direction after a press;it looks like the consequences of responses are selecting the responses that produce a goal result.

If this is what is actually going on (goal achievement via selection BY consequences) then there really is no difference (other than perspective and, possibly, predictive accuracy) between the reinforcement and PCT views of goal oriented behavior. Selection by consequences (reinforcement theory) produces purposeful behavior just as does selection of consequences (control theory).

I believe that the goal directed behavior of the reinforcement model is actually a lucky result of the geometry of the model's environment. My mistake was to imagine that the E. coli situation makes all consequences of action (bar presses) random; in fact, it does not. There is a systematic relationship (because of the environmental geometry) between a change in direction after a press and direction of movement before the press. When you eliminate this systematicity (which I did by having the dot randomly change its position as well as its direction after a press) the reinforcement model no longer works (the control model is, of course, unfazed).

My conclusion, based on these results, is that the success of the reinforcement model depends on a cooperative environment; in fact, the environmental contingency (in the E. coli case, the contingency between prior direction and consequent change in direction) is what makes the reinforcement model behave as it does.

The reinforcement model is not really a model of how the organism produces a particular goal result in a changing environment (ie. it is not a model of control); it is a model of what happens when an organism of a particular type interacts with an environment of a particular type.

Perhaps the best way to get this across is for me to try to develop a more "fool proof" version of the E. coli demo, one where all conceivable results of responses are random.

While I work on such a demo (which may not be possible to produce but what the hey) I think it would be a good idea to start modelling the operant ratio data with a reinforcement model. Maybe that's the best test of reinforcement theory after all -- see if it can account for the basic data that led to the development of the theory in the first place -- the data on "schedules of reinforcement".

Best Rick

Date: Wed, 14 Jun 1995 08:15:35 -0600
Subject: E. coli model

[From Bill Powers (950614.0630 MDT)]

Rick Marken (950613.2030)--

> Well, now that Bill Powers (950613.0815 MDT) has sold me down the river ;-),

Don't despair. There is no such thing as reinforcement -- you know it, I know it, and Bruce Abbott knows it. The environment does not have the power to act on an organism in such a way as to make it do any specific behavior. That it _appears_ to do so, however, is not disputable, just as no one disputes that the little man on the lawn ornament appears to be cranking the windmill and causing the wind to blow. We're talking about a situation much like the phlogiston-oxygen controversy, where everything hinged on the assumed direction of a basic effect. Was combustion caused by something that was emitted, or by something that was absorbed? Here, the question is equally basic: is the behavior of organisms controlled by events in the local environment, or are events in the local environment controlled by the behavior of organisms? As Bruce has said, we are talking about mutually-exclusive interpretations, but without added information of some kind we are in a Necker-Cube kind of situation. As long as information is missing, we can flip back and forth between the two views; each one, temporarily, seems convincing.

PCT brings in a lot of auxiliary information that makes the choice clear. If you disturb the controlled variable, the behavior changes -- for no apparent reason -- so as to oppose the effect of the disturbance. If the system's reference level changes, the same environment now appears to cause different behavior -- for no reason to be found in the environment. That is the route to disproving reinforcement theory: i.e., to showing what really controls what.

In the E. coli case, we haven't discussed on the net the effects of disturbances or of changes in goals. You did so in your experiment with multiple targets, but since your argument also contained a fallacy -- that there was no systematic effect of the tumbling on which to hang a reinforcement model -- you (we) left a place where your (our) thesis could be legitimately attacked, thus drawing attention away from the main point. Those whose life work would be negated if you are right can't be blamed for refusing to let the mistake go in order to see the reasonableness of the remainder of the argument.

> The subject's goal directed behavior in this task seems to be inconsistent with my qualitative understanding of reinforcement theory which says that responses (bar presses) are selected by their consequences (direction of movement after a press). Random consequences should select random responses.

The problem with the E. coli example is that the consequences are not totally random. If they were, control would probably be impossible (as Martin Taylor has occasionally remarked). In fact, because there are boundaries to the range of the random effects, when a consequence is near one boundary, a random change is more likely to take the consequence away from that boundary than toward it. That's the systematicity that allows E. coli to control (I think) -- and also what makes a reinforcement model possible.

So the E. coli demonstration is probably not the key to showing what is wrong with reinforcement theory, as we both once thought it was. We must go back to PCT basics: resistance to disturbance, and the arbitrariness of reference signal settings.

Best to all, Bill P.

Date: Wed, 14 Jun 1995 10:59:25 -0500
Subject: Gerry Spence, PCT Lawyer

[From Bruce Abbott (950614.1155 EST)]

If you have been tuning in on any of the television programs that review the day's happenings in the O. J. Simpson trial you probably have seen and heard Gerry Spence, the homespun, buckskin-jacketed Wyoming defense attorney who has never lost a criminal case. Gerry has just published a book entitled _How to Argue and Win Every Time_, and I figured, hey, this is just what I need for those CSG-L debates--I'll just read Gerry's book, and I'll never lose another argument! So yesterday I purchased the book and began to read it. It's an amazing book. According to Gerry, all you need do to win an argument is apply the fundamental insights of PCT.

I doubt that Gerry Spence has ever heard of PCT, but that's what it comes down to. From year of experience, Gerry has developed an intuitive understanding of human behavior that fits PCT like a pair of well-made buckskin gloves. I don't think he'd mind if I offer a quote from his book to support this claim:

> Understanding how power works: Power is first an idea, first a perception. _The power I face is always the power I perceive_. Let me say it

> differently. Their power is my perception of their power. Their power is _my_ thought. The source of their power is, therefore, in _my_ mind.
>
> The power others possess is the power I give them. Their power is _my gift_. I give them all the power in the universe, as, indeed, the faithful give to God, or I give them no power at all, as, indeed, is the quantum of power we too frequently allot to our children. If I have endowed the _Other_ with power that the _Other_ does not possess, then I face my own power, do I not? My own power has become my opponent, my enemy. On the other hand, if the _Other_ possesses power, but I do not perceive the _Other's_ power as effective against me, he has none --- none for me.
> (Spence, 1995, p. 33)

Now I've done it. You're going to go out and get your own copy of this book, you're going to read it, and then _you'll_ win all the arguments. Me and my big mouth. (;->

Regards, Bruce

Date: Wed, 14 Jun 1995 09:33:36 -0700
Subject: Love's Labours Lost

[From Rick Marken (950614.0930)]

Me:

> The subject's goal directed behavior in this [E. coli] task seems to be inconsistent with my qualitative understanding of reinforcement theory which says that responses (bar presses) are selected by their consequences (direction of movement after a press). Random consequences should select random responses.

Bill Powers (950614.0630 MDT) --

> The problem with the E. coli example is that the consequences are not totally random... So the E. coli demonstration is probably not the key to showing what is wrong with reinforcement theory, as we both once thought it was. We must go back to PCT basics: resistance to disturbance, and the arbitrariness of reference signal settings.

It is with some sadness that I admit that this is true. E. coli seemed like such a pretty little demo of the impossibility of a reinforcement model of purposeful behavior; and it is pretty as long as one looks only at direction of dot movement rather than _change_ in direction of dot movement as the consequence of responding.

> As Bruce has said, we are talking about mutually-exclusive interpretations, but without added information of some kind we are in a Necker-Cube kind of situation. As long as information is missing, we can flip back and forth between the two views; each one, temporarily, seems convincing.

Yes, and it seemed like the E. coli demo brought in that missing information; seeing that the consequences of responses were random would make it clear that these consequences could not be shaping the responses that produce a systematic result.

I loved my little E. coli demo. But I have to admit that, in its present form, it is not a rejection of the "control by consequences" view of the Necker Cube of purposeful behavior.

Back to the drawing board. Sob.:-(

Best Rick

Date: Wed, 14 Jun 1995 11:58:05 -0600
Subject: PCT Lawyer; Farewell to E. coli

[From Bill Powers (950614.1120 MDT)]

Bruce Abbott (950614.1155 EST) --

RE: Gerry Spence, PCT lawyer.

Nice find. I've been saying something similar for years, but on a different subject: The System. To people who rail against The System I say: _There is no System_. You have never seen it, you have never interacted with it: it does not exist, except in your own perceptions. All you have ever seen or interacted with is people, usually one person at a time. All the problems that seem to be given to you by The System are actually given to you by some specific person, for that person's reasons (which may, of course, include belief in The System). So you never have to deal with The System. All you have to do is deal with a person, then another person, then another person. That may be hard enough, but it is nowhere near as hard as dealing with a vast implacable all-powerful all-knowing System, a creation of your own mind.

Same principle as Spence's understanding of Power. Try it on Behaviorism, Cognitive Science, Physics, Government, Society -- and PCT.

Rick Marken (950614.0930) --

> I loved my little E. coli demo. But I have to admit that, in its present form, it is not a rejection of the "control by consequences" view of the Necker Cube of purposeful behavior.

> Back to the drawing board.

Yes, by all means back to the drawing board. E. coli failed because there was order in what we thought was randomness. There was a logical perception, which can be rendered as

{[(N2 > N1) AND (N1 > Ref)] OR [(N2 < N1) AND (N1 < Ref)]}
AND NOT
{[(N2 > N1) AND (N1 < Ref)] OR [(N2 < N1) AND (N1 > Ref)]},

which could be used as the basis for lengthening or shortening the delay before the next tumble. The fact that we had found a much simpler perception that accomplished the same result even more quickly, namely,

NOT (N1 < Ref),

made us think that there couldn't be any more complex perception that would work.

What we need is a simple task in which there is no natural boundary to the effects of a random change in output: where the next change in the controlled variable is not predictable from its present state. The easiest context would be a one-dimensional task. It may be that if we do find such a task, it will not allow control, either. But it's worth at least a bit of looking.

Best to all, Bill P.

Date: Wed, 14 Jun 1995 15:20:58 -0400
Subject: E.Coli experiment

[From Dag Forssell (950614 1220)]

I follow, appreciate and enjoy the discussion of reinforcement theory. As Rick expressed yesterday, I too recall that the probabilities of favorable / unfavorable tumbles were discussed and settled late last year. I do not quite recall how I visualized the spacial/probability discussion in my kindergarten physics mind, but if as Rick now suggests the distribution is an artifact of the geometry, how about this geometry:

Change the target to a line.

This will make it easier for E. Coli to reach in the physical world. It should be easy to turn into a version of the manual game TUMBLE on the demodisk. It brings symmetry to the question of "away from target" or "towards target" with the in between, parallel to target, direction being in the middle of a random distribution of directions, after a random tumble.

I trust this will be quite easy to model. Rick's awkward discontinuity of random change of location is avoided.

Best, Dag

Date: Wed, 14 Jun 1995 14:32:14 -0600
Subject: Love's Labours Recoverable?

[from Gary Cziko 950614.1920 GMT]

Rick Marken (950614.0930) pined:

> I loved my little E. coli demo. But I have to admit that, in its present form, it is not a rejection of the "control by consequences" view of the Necker Cube of purposeful behavior.

> Back to the drawing board. Sob.:-(

Your original E. coli demo may not have done what you wanted it to do, but I think your approach was a good one--finding the simplest example of goal-oriented behavior that could not be accounted for by reinforcement theory.

But do you really have to make it much more complicated to make the same point? Hasn't it already been agreed by Bruce Abbott that your modification which makes the E. coli instantly appear in a new location cannot be handled by his model?

And couldn't the same effect be achieved more realistically by adding a disturbance to the bacterium's movements (like a water current) that would push it while it was tumbling (with tumbling taking some time) while the food source remained anchored to one spot?

So perhaps making the E. coli model a little more complicated but also more realistic will accomplish the same goal as the too-simple model.--Gary

Date: Wed, 14 Jun 1995 15:43:34 -0500
Subject: A Different View

[From Bruce Abbott (950614.1510 EST)]

In my post of 950611.1635 EST I suggested something to the effect that ECOLI4a could be viewed as a form of adaptive controller something like the one Hans Blom has been describing. I've tried to make this view more visible in the diagrams below. The first diagram is that of an e. coli-type control system:

```
              dN/dt > 0
                 |     error
         ----->[  ]-------------->-----------------
        |                                          |
        |                                          |
        |                                          |
  dN/dt |                                          |
 [Input Function]                          [Output Function]
        |                                          |
        |                                        TUMBLE
        | N                                        |
        |                                      direction
        |                                          |
         -------------------[environment]----------
```

This is a one-way control system in which an error develops if the perceived nutrient rate of change is not positive. Unlike the standard e. coli control model, however, the error signal is a two-state, logical variable rather than a continuous one. If this variable is True (no error), the output function produces a tumble rate determined by PTS+; if it is false, the output function produces a tumble rate determined by PTS-. If PTS+ is greater than PTS- the result is negative feedback. In the ideal case, PTS+ = 0 and PTS- = 1.0. In that case, tumbling will never occur when dN/dt is positive and will always occur so long as dN/dt is not positive.

The second part of the system supplies the adaptation, and looks more or less as diagramed below:

```
                 dS/dt = 0
                     |     error
         --------->[  ]----------------
 dS/dt  |    dN/dt = 0                 |
        |         |   S+ or S-         |
        |    -- [  ]------->-------    |
        |   | S = dN/dt           |    V
    [Input Function]            [Parameters]--->[Output Function]
        |                            |                 |
        |                             ------<--------TUMBLE
        | N                                            |
        |                                          direction
        |                                              |
         -------------------[environment]--------------
```

This diagram isn't quite right, as I am not sure how to represent some of the relationships. The input function provides both the rate of nutrient change, S (dN/dt) and the change in nutrient rate, dS/dt. The change in nutrient rate will always be zero except following a tumble. If the is change is non-zero, one of the two output function parameters is adjusted. If S+ was present prior to the tumble (dN/dt positive), then PTS+ is changed; if S- was present (dN/dt negative) then PTS- is changed. The direction of these changes is determined by the error signal relating to dS/dt: if dS/dt is positive (improvement), P is incremented; if negative, P is decremented.

The problem I have with the diagram is in representing the fact that the value of dN/dt which applies when adjusting parameters is the value immediately prior to a tumble rather than its value at the time adjustment of parameters takes place. In effect there is a delay or lag that should be inserted between dN/dt and the connection to the parameter adjustment mechanism. This amounts to a brief storage or memory of the pre-tumble value. The diagram also does not make explicit the fact that dS/dt is computed across a tumble episode by comparing dN/dt before and after a tumble. I've added the line from TUMBLE to the parameter adjustment mechanism to suggest that parameter adjustment takes place only following each tumble, although I don't find this representation very satisfactory. Bill P., perhaps you could suggest a better way to construct this diagram?

Despite these problems, I think the diagram does highlight the parallel with Hans Blom's adaptive controller: feedback from the consequences of a tumble is used as a means of adjusting the low-level nutrient-rate control system's parameters so that control over nutrient rate is achieved.

As a suggestion, I think a "cleaner" example of this concept could be developed based on the standard e. coli control model, with the adaptation mechanism operating on the control system gain factor. To function properly, the adaptation control loop must be sluggish compared to the system it adjusts.

Regards, Bruce

Date: Wed, 14 Jun 1995 18:37:49 -0500
Subject: Final Scene at the Long Branch

[From Bruce Abbott (950614.1835 EST)]

Scene IV: The Ranch House

It has now been over three hours since Pa rode off to town to confront Black Bart, and Ma and the Kid are gettin' mighty worried. "Kid," Ma says, "this doesn't look good. Pa should have been back by now. Head over to the Long Branch and find out what happened. And Kid ..."

"Yeah, Ma?"

"Be careful."

Scene V: Back at the Long Branch

The Kid pulls up to the Long Branch and throws the reigns of his horse over the hitching rail. The sounds of the honky-tonk piano and the cacophony of many simultaneous conversations drift out through the saloon doors. Fearing the worst, the Kid slowly pushes against the swinging doors and walks inside. Everything looks normal, no sign of anything amiss. Suddenly a familiar voice calls out over the din. Its Pa!

"Hey Kid, com'on over an' have a seat!"

There, sitting at one of the Long Branch'es round oak tables, are Bill McGraw and Black Bart. The table is cluttered with empty glasses and peanut shells. Neither of the two men look particularly dead, or even slightly wounded. In fact, they seem to be having a pretty good time. His mind reeling in disbelief, the Kid walks to the table and plunks down in an empty chair.

"Kid," says Bill, "the whole thing's been a misunderstanding. Once we got it all sorted out, there wasn't nothin' left to argue about. Have a drink. Me an' Bart here were just gettin' warmed up."

>Bill Powers (950613.0815 MDT) --

Thanks, Bill, for struggling with the problem I presented you, for being willing to keep reexamining and testing your (and my) understanding of the mechanism by which the ECOLI4a program was able to do what it does. It would have been much easier just to write my claims off as the misperceptions of a sadly confused psychologist still trying to preserve some vestige of belief in the old theory of behavior. There were times when you had me seriously worried that perhaps your interpretation was right, that I was missing something critical that you were seeing plainly. That forced me to critically reexamine my own understanding of the program, and as a result I gained not only a greater faith in my analysis but was able to develop a quantitative analysis of the program's performance. The debate was not a simple assertion of claim and counter-claim, but a true attempt by both sides of the dispute to find the basis of their disagreement and resolve it. It is really only through such testing that real understanding can emerge. If this happens, everyone wins. Whose argument prevailed isn't all that important.

When the resolution in this case turned out not to favor your interpretation, you did not hesitate to say so. There's a name for people who willingly change their opinions on the basis of sound evidence and argument. We call them scientists.

> Come on, somebody else make a comment that is really truly wrong. I need someone to beat up on.

Give me a post or two, after I've gotten over the shock of being right about something for once... (;->

Regards, Bruce

Date: Wed, 14 Jun 1995 21:50:23 -0600
Subject: Re: Reinforcement model

[From Bill Powers (950614.1812 MDT)]

Bruce Abbott (950614.1510 EST) --

We have found out why the reinforcement model works with the E. coli situation. The question still remains whether there is a control task (doable by a human being) that can't be handled by the principle of reinforcement + discriminative stimulus. If there is one, it could serve to distinguish these theories in a direct way.

The basic problem is as follows:

Given:

Q = quantity to be controlled.
Q1 = quantity before tumble, Q2 = quantity after tumble

```
  R+ = Q2 > Q1,  R- = Q2 < Q1
  S+ = Q1 > 0,   S- = Q1 < 0.
```

If we plot these conditions on a line with maximum positive (+1) at the right and maximum negative (-1) at the left, with zero in the center, we get

```
                                |
(1)                             |     Q1                Q2
(2)   Q2           Q1           |
                                |
-1------------------------------0------------------------------+1
                                |
(3)                       Q1    |    Q2
(4)                       Q2    |    Q1
                                |
```

The upper two cases are the "wrong" cases which work against going up the gradient:

Line (1) (Q1 > 0) AND (Q2 > Q1) inc(PTS+)
Line (2) (Q1 < 0) AND (Q2 < Q1) dec(PTS-)

The lower two cases are the "right" cases:

Line (3) (Q1 < 0) AND (Q2 > Q1) inc(PTS-)
Line (4) (Q1 > 0) AND (Q2 < Q1) dec(PTS+)

The reason that the wrong cases occur only 25% of the time can easily be seen. In line (1), the average value of Q1 (which has to be greater than zero) will be +0.5. Q1 has a 50% chance of being greater than 0. Q2, which can be anywhere in the range from +1 to -1, has a 25% chance of being greater than the average Q1. So the total probability of this case is 0.5*0.25 or 12.5%. The same applies to the second line.

In line 3, it is required that Q1 (average value 0.5) be less than zero (50% chance) and that Q2 be greater than the average Q1 (75% chance). 0.5*0.75 = 37.5%. The same applies in line 4.

This suggests that to construct a task in which a reinforcement model can't work, we must equalize the "good" probabilities and the "bad" ones (or make the imbalance go the other way). That is, whenever a tumble occurs, it must have a 50% probability of making Q larger regardless of the value of Q, present or past, and also regardless of whether the next value of Q is larger or smaller than the present value.

It will take some thinking to devise such a task where the output is still a random change in (whatever) that affects Q. Perhaps you can help -- or help to prove that such a task is impossible.

Your diagram can be made to look simpler if we just use the before-after values of Sx (where x means + or -). The value S2 - S1 or R is, roughly speaking, dS/Dt.

```
                              enable PTS adjustments
                    ----------------<-------------------
                   |            on next iteration        |
                   V                                     |
              ----------                                 |
      Rx --->|          |-- R+&S+ -->--------------      |
      |      |          |                          |     |
    dS/dt    | MINTERM  |-- R-&S+ -->-----------   |     |
      |      |          |                       |  |     |
    [d/dt]   | EXPANDER |-- R+&S- -->-----      |  |     |
      |      |          |             -  +|    -| +|     |
      Sx --->|          |-- R-&S- -->-->[PTS-] [PTS+]    |
      | \     ----------                  |      |       |
      |   -------------------------->  Delay1  Delay2    |
      |          S+ -> Delay1            \     /         |
      |          S- -> Delay2             TUMBLE---->----
   S=dNut/dt                                |
      |                                 direction
      |                                     |
       ---<----------[environment]-<--------
```

I'm not sure that's any improvement -- I didn't get the error signal in, and it really should enter right where Sx is. S+ really means S > Ref. Also I probably got the signs wrong for the PTSs.

\>   Despite these problems, I think the diagram does highlight the parallel with Hans Blom's adaptive controller: feedback from the consequences of a tumble is used as a means of adjusting the low- level nutrient-rate control system's parameters so that control over nutrient rate is achieved.

Right. In general, reorganization or adaptation means adjusting parameters on the basis of some measure of system performance. But see note at end.

The reinforcement model has one deficiency, which is that it will adapt only to attractants. In order to make it adapt as appropriate for an attractant or a repellent, something would have to detect the kind of input that is being sensed, and change the definition of S+. You can manually change the definitions, but that's not "adaptation." As it stands, the model doesn't know whether S+ is a good thing or a bad thing. It assumes it's good. If you say that a positive dNut is to be avoided, you have to redo the model, don't you?.

> As a suggestion, I think a "cleaner" example of this concept could be developed based on the standard e. coli control model, with the adaptation mechanism operating on the control system gain factor. To function properly, the adaptation control loop must be sluggish compared to the system it adjusts.

If I were designing an adaptive E. coli system from scratch, I'd use a two-level control system and not the reinforcement model. The reinforcement model, although it works, is far from the simplest model that would work. It isn't really necessary to use information about the change in S across a tumble. There's probably some mathematical manipulation that would show that R is redundant. The fact that we can get the same result in a simpler way, without using R, strongly suggests that the redundancy could be proven, but I'm not up to proving it.

Come to think of it, your approach isn't really like Hans' model, because there's no world-model in it -- no model of the link between the output action and the controlled variable Q. If you cut off the input to your model, it would not go on producing the same output as before.

I'm not sure we should even call the reinforcement model, as it stands, an "adaptive" model, because it can produce only the one "adaptation": PTS+ -> 0 and PTS- -> 1. If it could automatically detect the difference between repellents and attractants it would be properly adaptive.

Best, Bill P.

Date: Fri, 16 Jun 1995 08:54:21 -0500
Subject: Reinforcement Model

[From Bruce Abbott (950616.0850 EST)]

>Bill Powers (950615.1230 MDT)

> This is the opposite of what I meant by "manually" changing the model: you're "mentally" changing it by changing what you mean by S+, without making any corresponding change in the model.

Yes, the point was that S+ and S- are NOT what gets "manually" changed to make the model work with repellants. Perhaps I misunderstood, but you seemed to be saying they were:

>> In order to make it adapt as appropriate for an attractant or a repellent, something would have to detect the kind of input that is being sensed, and change the definition of S+. You can manually change the definitions, but that's not "adaptation."

I needed to address that point before I could go on to state what _does_ get changed.

> OK, after all that talk you ended up realizing that you have to reprogram the model to make it work with a toxin.

No, I had this in mind from the start. But I had to get Point 1 out of the way before proceeding to Point 2.

> Your suggestions about how to do this are reasonable. When you implement them, you will have a truly adaptive model, if not the simplest one possible.

Good, we agree.

> So is learning NOT "behavior determined by the internal structure of the organism and the properties of its environment?" When we model anything that resembles learning, we come down to specialized circuits that have the effect of modifying parameters in other circuits. What else could learning be?

Of course it is. But not all "behavior determined by the internal structure of the organism and the properties of its environment" involves learning, which is a very different point and the one I'm trying to make here.

> The notion of "hard-wired" isn't very useful: if you hard-wire multipliers into the system, it can change its organization just as readily as it could do under "software" control. Anything you can do with a program can be done with hard-wired components. If you stop and think about it, anything you can do with a program is ALWAYS done with hard-wired components.

Yes, but consider the control model for e. coli. The model has a certain structure, and so does the environment. Throughout the simulation the model's structure never changes. Either it controls or it doesn't. That's what I mean by "hard-wired": the structure is in place (including parameter values) and is not modified _on the basis of experience_.

A reinforcement model, in contrast, is built on the assumption that experience is a crucial factor determining how the organism will behave in a given environment. Changes take place in the structure (system parameters), owing to the consequences of the organism's own behavior, and these changes alter the way the organism will behave in that environment in the future. Steady-state behavior emerges from these changes acting in combination with current environmental conditions.

Thus, if there is no basis for learning (e.g., responses have no systematic consequences), there is no reason to expect that a learning-based model should apply.

> If we remove, one at a time, all the examples in which control theory works as well as or better than reinforcement theory, sooner or later there will be no pure examples of reinforcement left. That is, there will be no case in which ONLY reinforcement theory can explain the data.

That sounds like one long, tedious research project to me. Be my guest.

> It may be true that "the failure of the reinforcement model in this situation would say nothing at all about its validity", but that is also true of SUCCESS of the reinforcement model. All that success shows is that the reinforcement model fits the observations, not that it explains them correctly in terms of the actual mechanisms involved.

True enough, but the same can be said for any model (see Popper).

> Wherever both PCT and reinforcement theory can explain the same observations, choosing between the models has to depend on something other than the fit of predictions to data.

Yes, such as parsimony and "elegance." Initially those were the only criteria that favored Copernicus over Ptolemy. But eventually other, more direct ways to test the two views were developed, based on structural differences in the models which led to different implications (e.g., crystalline spheres were inconsistent with the orbits of comets). The same will be true, I think, for PCT versus reinforcement theory.

> From what Bruce says, if reinforcement theory doesn't explain how we learn to do the task in your new program, that will only show that reinforcement theory isn't _expected_ to work in that situation.

Not so fast. If you can show that behavior is maintained under conditions in which the consequences of behavior are unsystematic (i.e., there is no consistent reinforcement or punishment of responses), you have demonstrated that reinforcement is not a _necessary_ condition for behavior. But is it a _sufficient_ condition? Your experiment won't say.

A serious problem for PCT in its battle against reinforcement theory is that PCT as currently developed does not offer a really well-developed explanation for how the observed performance under given conditions is acquired. Where do all these control systems come from? How do they get established? Reinforcement theory describes the process of acquisition and then assumes that the factors leading to acquisition continue to maintain the behavior in the steady-state. What could be simpler? This is what you're competing against.

Regards, Bruce

Date: Fri, 16 Jun 1995 09:57:15 -0500
Subject: "r." coli

[From Bruce Abbott (950616.0955 EST)]

>Rick Marken (950615.0900)

> Here is a new version of the E. coli task that I will call R. coli because (I think) all consequences of responses are completely _R_andom.

> The task is (surprise) to keep a cursor aligned with a target. The subject affects the cursor by pressing the mouse button (the response). The consequence of this response is that the cursor moves to one of five randomly determined positions on the screen. One of these five positions is the target position.

> So the result of any response is one of five randomly determined cursor positions. The cursor remains in this randomly determined position until another response occurs or until the position is changed (randomly) by an external disturbance.

> I believe (but, given my experience with E. coli I am not certain) that there is no way for a reinforcement model to explain the results of this R. coli experiment. All consequences of responding seem (to me) to be completely random; there seems to be no systematicity . But I'm the one who wants to show that reinforcement theory cannot explain control so I may not be looking carefully enough. So I hope Bruce Abbott will take a shot at developing a reinforcement model of R. coli.

This sounds like a variable-ratio schedule of reinforcement to me, Rick. Let's call "being on-target" the reinforcer. Pressing the mouse button is the response. Cursor movement is equivalent to a "feedback" beep that tells you the button is working (like the click that is heard when a lever is depressed far enough to operate the switch).

Situation 1: Cursor not on target.

Pressing the mouse button moves the target. Responses are emitted until the cursor moves to target or the behavior extinguishes from lack of reinforcement (whichever comes first). Occasional success reinforces button-pressing, thus maintaining the behavior. Success is programmed on a VR-5 schedule (on average, one in 5 responses will be reinforced, but the actual number between two reinforcements varies randomly from 1 to some upper limit.

Situation 2: Cursor on target.

You're "at the food-cup": no need to press the lever to obtain reinforcement, because you've got it.

Disturbances remove you from the reinforcement condition, taking you back to Situation 1 (just as eating the food in the cup takes the rat back to the food-absent situation, necessitating further lever-pressing).

So much for r. coli. NEXT!

> Don't worry Bruce; if reinforcement theory can handle this one I'll go right back to the drawing board again. I'm a glutton for punishment (and reinforcement).

I guess so. I'd keep the drawing board out. (;->

Regards, Bruce

Date: Fri, 16 Jun 1995 10:25:59 -0500
Subject: Correction

[From Bruce Abbott (950616.1025 EST)]

It never fails! As soon as I press the "send" button the mistakes jump off the page at me:

>Bruce Abbott (950616.0955 EST)

> Situation 1: Cursor not on target.

> Pressing the mouse button moves the target. Responses are emitted until the cursor moves to target or the behavior extinguishes from lack of reinforcement (whichever comes first). Occasional success reinforces button-pressing, thus maintaining the behavior. Success is programmed on a VR-5 schedule (on average, one in 5 responses will be reinforced, but the actual number between two reinforcements varies randomly from 1 to some upper limit.

The first sentence should read "Pressing the mouse button moves the cursor." Also, I left out the close-paren in the last sentence.

ADDENDUM

Rick suggested that having the cursor move to a non-target following a response would count as punishment, so I should explain why it doesn't. To act as a punisher, a consequence must make things worse than before the response. What you have here is "not on target" before the response and "not on target" after the response. No change in situation, no punishment.

Regards, Bruce

Date: Fri, 16 Jun 1995 09:42:58 -0700
Subject: Reinforcement: Neither necessary nor sufficient

[From Rick Marken (950616.0945)]

Me:

> I believe (but, given my experience with E. coli I am not certain) that there is no way for a reinforcement model to explain the results of this R. coli experiment. All consequences of responding seem (to me) to be completely random; there seems to be no systematicity . But I'm the one who wants to show that reinforcement theory cannot explain control so I may not be looking carefully enough. So I hope Bruce Abbott will take a shot at developing a reinforcement model of R. coli.

Bruce Abbott (950616.0955 EST) --

> This sounds like a variable-ratio schedule of reinforcement to me, Rick.

[Verbal explanation of R. coli]

> So much for r. coli. NEXT!

That was a very nice story about VR schedules and all; you do an excellent B. F. Skinner imitation. But I know that you are a better scientist than that.

What I want to see is a working reinforcement model of R. coli (just like the working reinforcement model you developed of E. coli).

After having analyzed my R. coli demo a bit more myself I think I can say with even greater confidence that there is no way a reinforcement model can account for the subject's behavior in this task.

I claim that the R. coli demo shows that reinforcement theory is neither necessary NOR sufficient as a model of control. I believe that there is no way to change the probability of response based on the consequences of responses in this task so that responses keep the cursor on target as well as the subject does.

R. coli exposes the illusion of reinforcement -- I think. The only way you can show that it does expose this illusion is by presenting a working reinforcement model of the behavior in the R. coli task.

Best Rick

Date: Fri, 16 Jun 1995 09:54:34 -0700
Subject: I correct, therefore I control

[From Rick Marken (950618.0955)]

I said:

> The only way you can show that it does expose this illusion is by presenting a working reinforcement model of the behavior in the R. coli task.

I meant:

The only way you can show that it does NOT expose this illusion is by presenting a working reinforcement model of the behavior in the R. coli task.

Best Rick

Date: Fri, 16 Jun 1995 14:42:22 -0500
Subject: Proving the obvious--the hard way

[From Bruce Abbott (950616.1440 EST)]

>Rick Marken (950616.0945) --

> That was a very nice story about VR schedules and all; you do an excellent B. F. Skinner imitation. But I know that you are a better scientist than that. What I want to see is a working reinforcement model of R. coli (just like the working reinforcement model you developed of E. coli).

Well, thank's for the compliment, but I decline. In this very simple application, if logic is not enough, a model won't do it either. You're asking me to do a lot of work to prove the obvious, and when I do, you will simply develop yet another "challenge" and it'll be Round Three. I've got better things to do.

Tell you what. YOU write the reinforcement model: I've described its logic well enough that you should be able to implement it. Post the code and I'll tell you whether it conforms to my little "story." Fair enough? (;->

By the way, I'm going on a brief vacation starting Saturday morning, and won't be back until Wednesday at the earliest. That reinforcement model will be waiting for me when I get back, right?

Enjoy the rest! Regards, Bruce

Date: Fri, 16 Jun 1995 13:57:14 -0700
Subject: It's getting more obvious

[From Rick Marken (950616.1400)]

Bruce Abbott (950616.1440 EST) --

> In this very simple application, if logic is not enough, a model won't do it either.

I think you may be right again.

> Tell you what. YOU write the reinforcement model: I've described its logic well enough that you should be able to implement it. Post the code and I'll tell you whether it conforms to my little "story." Fair enough? (;-

I've already tried one simple version of the reinforcement model and it seems to work alright. The reinforcement model that worked for E. coli doesn't work for R. coli, but it looks like some kind of control by consequences model will work.

> That reinforcement model will be waiting for me when I get back, right?

I'll look at it more closely but it's beginning to look like you're right again. Maybe this "random consequences" approach to dealing with reinforcement theory isn't the way to go.

But I am surprised by one thing. You don't seem to be working very hard at thinking of ways of showing what's wrong with reinforcement theory. You seem to know that the direction I'm taking (developing control tasks where the consequences of responding are random) is a dead end. It would be nice if, besides showing how reinforcement theory can deal with my "random consequences" demos with no problem at all (though I still have hope), you could spend some time suggesting ways to rule out the reinforcement theory explanation of control.

Maybe on your vacation you could try to think up some ways to show that consequences don't actually strengthen the responses that produce them. Then I wouldn't feel so bad about spending the same time developing the models that show that it can sure look like they can.

Best Rick

Date: Fri, 16 Jun 1995 17:21:23 -0500
Subject: Keepin' Busy

[From Bruce Abbott (950616.1720 EST)]

>Rick Marken (950616.1400)

> I think you may be right again.

Rick, I'm ALWAYS right. (Except when I'm wrong, which is fairly often.)

> But I am surprised by one thing. You don't seem to be working very hard at thinking of ways of showing what's wrong with reinforcement theory.

Golly gee Boss, don't a feller git some time off fer good behavior? You've kep me so busy workin' hard a showin' how reinforcement theory CAN account for certain data that I've hardly had time left for anythin' else. Now ya want me ta work on my VACATION? What a slave driver. An' another thing, you pay lousy. I keep waitin' fer th' check ta come in, but I ain't seen nothin' since I JOINED this outfit! When the heck is PAYDAY?

O.K., I'll think about it. Th' check is in the mail, right? (;->

Regards, Bruce

Date: Fri, 16 Jun 1995 18:43:06 -0600
Subject: Re: Reinforcement model

[From Bill Powers (950616.1130 MDT)]

Bruce Abbott (950616.0850 EST) --

We're in agreement; just a little language difficulty. When I said we have to change the definition of S+, I meant that we have to change the relationship of S+ to the physical situation, not that we have to rename the physical situation or S+.

> Yes, but consider the control model for e. coli. The model has a certain structure, and so does the environment. Throughout the simulation the model's structure never changes. Either it controls or it doesn't. That's what I mean by "hard-wired": the structure is in place (including parameter values) and is not modified _on the basis of experience_.

This is a fuzzy distinction; we draw the line where convenient. But I agree that it is convenient to draw the line.

> Changes take place in the structure (system parameters), owing to the consequences of the organism's own behavior, and these changes alter the way the organism will behave in that environment in the future.

What you said was

Behavior --> consequences --> structural changes --> behavior etc.

This makes the structural changes depend directly on the consequences, as if there were a predetermined link between the consequences and the changes in structure. So this gives the environment a way of controlling behavior, through altering the structures that determine the relationship between behaviors and consequences.

But now consider this way of looking at it:

```
Behavior -> consequences
                        \
                         \
                        error -> structure -> behavior --> etc.
                       /  |            ^
                     /    v            |
desired consequences      mechanism for
                      converting error into
                       change in structure
```

This picture says that there is a control loop in which the consequences of behavior act through the EXISTING structure to create behavior that affects the consequences. This is the basic control loop, with no special effect of consequences on structure. However, inside the organism there is a mechanism that alters the structure until the error is minimized. This is another control loop which is concerned with the quality of control (minimum error) rather than the state of the controlled variable. Without this second loop, the structure would not change, and behavior would remain in a fixed relationship to its consequences.

There is a big difference between the first and second models. In the first model, structural changes are due to an effect of a consequence on nervous system structure. In the second model, a consequence has no effect at all on structure; it is simply an input variable that is controlled as usual. What changes structure is a control process inside the behaving system. Also, in the second model there is a reference signal that determines what state of the consequence will constitute "no error", and thus which states of the consequence will result in structural changes, and in which direction.

These are the basic differences between reinforcement theory and the learning theory appropriate for PCT. Under reinforcement theory, something about the consequence of behavior causes structural changes in the organism that result in behavior either increasing or decreasing. Under PCT, the consequences of behavior have no ability to change the organism's structure; instead, the organism itself initiates structural changes in its own subsystems as a means of gaining control over the consequence.

The final appearance is the same under either theory -- as long as reference signals in the organism do not change, and as long as there are no independent disturbances of the controlled consequence. If the reference signal changes, however, we will find that a consequence of behavior which formerly apparently caused an increase will now cause a decrease in behavior, or vice versa. If there is a disturbance of the consequence, we will find that behavior changes to restore the consequence to its previous condition: the same consequence is now produced by a different behavior. Reinforcement theory is not capable of explaining either of these types of change. Or, remembering the discussions of the past few months, so I contend.

The logic of reinforcement theory, when cast in terms of S+ and R+ and so forth, sounds formal and precise. However, in plain language it sounds simpler.

What we observe is this:
There is some indicator that occurs in a situation where behaving in a certain way will result in an increase in Q. The frequency with which behavior appears (when the indicator is present) increases.

The reinforcement explanation is this:
The increase in Q reinforces the effect of the indicator on the behavior that produces Q. What this means is that Q changes the link between the indicator and the behavior so the behavior becomes more likely when the indicator is present.

What we observe, to expand, is that

1. Q+ and S+ --> increased frequency of B | S+
2. Q- and S+ --> decreased frequency of B | S+
3. Q+ and S- --> increased frequency of B | S-
4. Q- and S- --> decreased frequency of B | S-

Note that these are descriptions of what we observe under the various conditions.

Reinforcement theory offers this explanation:

1. Q+ and S+ --> increase the probability of B | S+
2. Q- and S+ --> decrease the probability of B | S+
3. Q+ and S- --> increase the probability of B | S-
4. Q- and S- --> increase the probability of B | S-

The only thing that the explanation adds to the observations is that an increased frequency of B|S is due to an increased probability of B|S. Since the only way we have to observe probabilities is to observe frequencies of occurrence, the explanation is essentially just a repetition of the observation.

The whole question for modeling is HOW the conditions on the right are made to follow from the conditions on the left. If we just set up a logical computation of B|S on the basis of the states of Q and S, we don't have a model; we're just re-expressing the observations as a computation. We haven't proposed any actual _mechanism_.

> A serious problem for PCT in its battle against reinforcement theory is that PCT as currently developed does not offer a really well-developed explanation for how the observed performance under given conditions is acquired. Where do all these control systems come from?

As I've been trying to show above, reinforcement theory doesn't offer a well-developed explanation, either, although it does describe a particular orderly way in which some kinds of learning might take place. Everybody knows that somehow the kinds of learning we see require that new behavioral organizations appear and that old ones be modified. But so far, we have few specific models (like the Extended Kalman Filter model and the E. coli reorganization model) that come anywhere near being a _mechanism_ of learning. It's not enough to say that if there are signs in the environment that behaving in a certain way will succeed, organisms presented with those signs will change their behavior in that way. If we want a real model, we have to make guesses about HOW they manage to recognize the information, and HOW they are able to alter their own structures in the right way.

> Reinforcement theory describes the process of acquisition and then assumes that the factors leading to acquisition continue to maintain the behavior in the steady-state. What could be simpler? This is what you're competing against.

It's hard to compete against because it sounds clear and simple and actually says nothing useful. The whole question is WHAT are the factors that maintain behavior in the steady state, and HOW do those factors maintain it? The reinforcement explanation is simply glib, and those who are impressed by it are not likely to see why it is empty.

Best, Bill P.

Date: Sat, 17 Jun 1995 04:29:54 -0600
Subject: The Power of Reinforcement Theory

[From Bill Powers (950617.IHateToSay)]

Bruce Abbott (950616.1230) --

Some added thoughts after sufficient reorganizing time has passed.

> Reinforcement theory describes the process of acquisition and then assumes that the factors leading to acquisition continue to maintain the behavior in the steady-state. What could be simpler? This is what you're competing against.

Actually, I think I'm beginning to get the hang of explaining behavior with the powerful competing method of reinforcement theory. Tell me if I'm getting this right.

Reinforcer = R
Discriminative stimulus = S
Behaviors = Bn

----

1. Learning to call a bottle "bah"

R+: mother smiles
R-: mother frowns
S+: bottle appears
S-: ball appears
B1: utter "bah", B2 = utter "dih", B3 = cry

If mother smiles only when the ball appears, the probability of uttering "bah" when the ball appears increases. Eventually, every time the ball appears, "bah" will be uttered.

----

2. Learning to move a cursor to a target.

R+: Finger-target separation decreases
R-: Finger-target separation increases

```
S+:  Finger left of target
S-:  Finger right of target
B:   Move finger left, right

If separation decreases
 if finger left of target
  increase Pr{ move finger right }
 else if finger right of target
  increase Pr{ move finger left }

If separation increases
 if finger left of target
  decrease Pr{ move finger left }
 else if finger right of target
  decrease Pr{ move finger right }
```

Eventually, when the finger is to the left of the target, the probability of moving the finger to the right will be 1, and when the finger is to the right of the target the probability of moving it left will be 1. When the finger is neither left nor right, there will be no discriminative stimulus and the finger will move neither left nor right.

Hey, this is a lot simpler than control theory! Let's get more ambitious.

R: remaining balanced and standing upright on deck of ship.

S: waves tilt ship to North, South, East or West

B: lean more to North, South, East, or West relative to deck

I don't need to repeat the analysis: eventually a tilt to the North will result in a lean to the South, and so forth around the compass.

It seems pretty obvious that reinforcement theory can explain the learning of any behavior. Let's go all-out:

R: Mathematicians say you're getting closer to or farther from proving Fermat's Last Theorem.

S: Mathematicians say current step of proof is correct or incorrect

B: execute any of a list of legal next steps

```
If mathematicians say you are getting closer to the proof
  If they say step of proof is correct
    increase Pr{ executing that step of the proof }
  else if they say step of proof is incorrect
    increase Pr{ changing that step of the proof }
else If mathematicians say you are getting farther from the proof
  if they say step of proof is correct
    decrease Pr{ executing that step of the proof }
  else if they say step of proof is incorrect
    decrease Pr{ changing that step of the proof }
```

So eventually you will cease to execute steps that lead mathematicians to say you are farther from the proof, and will execute only those steps that lead mathematicians to say you are closer to the proof. When mathematicians cease to say that a step is either correct or incorrect, there will be no discriminative stimulus and no further action. This explains how the proof of Fermat's Last Theorem was completed.

Not only is reinforcement theory powerful enough to explain the learning of any behavior whatsoever, it isn't even necessary for one to understand anything about brain function, physiology, physics, chemistry, control theory, mathematics, linguistics, and so forth. Reinforcement theory is the Grand Master Theory of behavior; all else is mere detail.

Best, Bill P.

Date: Sun, 18 Jun 1995 12:23:39 -0400
Subject: PCT & Reinforcement

[FROM: Dennis Delprato (950618)]

>Rick Marken (950616.1400)

>>Bruce Abbott (950616)

>> (Various comments on 'reinforcement model')

> I've already tried one simple version of the reinforcement model and it seems to work alright. The reinforcement model that worked for E. coli doesn't work for R. coli, but it looks like some kind of control by consequences model will work.

> I'll look at it more closely but it's beginning to look like you're right again. Maybe this "random consequences" approach to dealing with reinforcement theory isn't the way to go.

For some time I've had the 'feeling' that the selection by consequences description of how reinforcement 'works' will not be easy to displace, and not because this theory is of some sort of ultimate value. Rather, 'selection by consequences' is a type of radically descriptive account of _procedures_ and outcomes. As such it is of great value in that previous attempts to provide general descriptions of 'reinforcement' asked us to accept this or that hypothetical construct, e.g., strengthened S-R bonds, response strengthening, confirmed expectations.

Rick's attempts to use PCT models to test the generality of the selection by consequences account may be doomed to support selection by consequences basically because thus far he has operated according to selection by consequences's terms. That is, his models seemed to have been based on _their_ fundamental units-- R - C and S - R - C (R = operant response, C = consequence, S = discriminative stimulus). The units are built into the reinforcement procedures and thus have inherent in the models because the models were based on the procedures.

Another feature of selection by consequences theory in addition to its radically descriptive nature makes it difficult to overthrow. This is the fact that reinforcers are functionally defined. If particular events follow occurrences of operant responses and the rate of the operant does not change, one has not established the operation of reinforcement. One implication of this for a PCT modeler seems to be that if one sets out beforehand to model reinforcement procedures, the model will always be compatible with a reinforcement description.

If PCT offers a new fundamental unit of psychological behavior, then PCT is going "underneath" the operations of reinforcement. It is getting at what is more generally going on than what we observe on the surface. The question, then, is how to go beyond the surface of selection by consequences reinforcement. One place to look might very well be the work on feedback functions. It seems to me that the molar behaviorists (where one finds feedback functions) have departed from Skinnerian "molecular" selection by consequences. They say, I think, "When responses are modified by contiguous relationships between responses and consequences, it is not what one sees directly (contiguous relationships) but what one does not see directly that is more generally important for describing what is going on."

Dennis Delprato

Date: Sun, 18 Jun 1995 15:06:55 -0600
Subject: Then a miracle occurs

[From Bill Powers (950618.1430 MDT)]

Dennis Delprato (950618) --

> If PCT offers a new fundamental unit of psychological behavior, then PCT is going "underneath" the operations of reinforcement. It is getting at what is more generally going on than what we observe on the surface. The question, then, is how to go beyond the surface of selection by consequences reinforcement.

What has me climbing the wall is the use of the term "selection by consequences" when to me it is obvious that nothing of the sort is going on.

I've been watching the U.S. Open golf tournament on TV. Shinnecock is a tough course and club selection for short approach shots is critical. If the club is too lofted the wind makes the ball drift, and if not lofted enough the ball runs when it hits the hard green as well as falling short. So what selects the club? The consequence of drifting in the wind or falling short, or THE PLAYER? I contend that it is the player who does the selecting, using information about previous experience to be sure, but still using a process of selection that lies in the player and not in the ball.

The behavior of a golf ball is just not the sort of thing that can "select." Selection of behavior is a process of weighing many factors and choosing an action that is estimated -- by a system capable of making estimates -- to produce a desired result. The ball has no preference for how it will behave and mechanism for making estimates or choosing on that basis; only the player can have such a preference, or draw conclusions from present-time data.

> One place to look might very well be the work on feedback functions. It seems to me that the molar behaviorists (where one finds feedback functions) have departed from Skinnerian "molecular" selection by consequences. They say, I think, "When responses are modified by contiguous relationships between responses and consequences, it is not what one sees directly (contiguous relationships) but what one does not see directly that is more generally important for describing what is going on."

Left out of this way of putting it is the actual mechanism that does the modifying. That mechanism has to lie within the organism -- where else could it possibly be? The mere fact that a response has a certain consequence can't modify anything else. It's just a fact, something that happened. I don't think that proposing that contiguous relationships can modify responses is any better than saying that consequences can select behavior. Neither one makes any sense to me. How on earth could a contiguous relationship do anything but be a contiguous relationship? Wherein lies its ability to modify a response?

Help me out, Dennis. Am I the only one in the world who sees statements like this as invoking magic? Here is a consequence; here is a contiguous relationship. Then a miracle occurs, and responses are modified or behavior is selected. Doesn't anyone else see something vital missing from these statements?

Best, Bill P.

Date: Sun, 18 Jun 1995 22:03:07 -0400
Subject: Re: PCT & Reinforcement

<[Bill Leach 950618.16:38]
>[Dennis Delprato (950618)]

Dennis; I think that you have "hit upon" something that is even more significant than first might appear...

EAB work appears (to me anyway) as a genuine effort to conduct scientific research in a manner consistent with the principles of scientific methodology.

EABers apparently DO ask serious questions concerning the validity of their methods _and then_ test their conclusions. Furthermore, they repeat their experiments and create scripts of sufficient detail to allow others to do so also.

EAB is likely wrong, in that they are always (nearly always?) setting up an environment that produces disturbances to intrinsic references. It is highly likely that the intrinsic reference set for any particular type of being will be uniform across the species.

The EAB view then is "mislead" by the perceived consistency between reinforcers and observed behavior for at least two reasons:

The first is the aforementioned disturbance to intrinsics and the second is that much of the actual behavior of the test subjects is ignored (understandable but still a significant source for misunderstanding what is actually going on).

A well conducted EAB experiment then _should be_ "consistent" with a PCT explanation for the same phenomenon. That is a well conducted EAB experiment will be an accurate _description_ of exactly what happened.

The EAB conclusions will, from a PCT point of view of course, be incorrect. Indeed, I suspect that EAB research experience provides its' own ample evidence of "problems" with the explanations. That is, changes to the experimental parameters do not always lead to the predicted observed behavior (remembering also that some of the significant behavior may just plain "not be noticed").

Nonetheless, it just might be EAB where PCT could eventually find its' most ardent supporters as EABers begin to see that when you "flip the cube" the new dimensions are aligned with what used to be "behavioral anomalies".

-bill

Date: Mon, 19 Jun 1995 09:03:01 -0400
Subject: Re: Then a miracle occurs

[Dan Miller (950619)]

Bill Powers (950618)

Very nice post. The invocation of magic is everywhere around us. I have to work hard to keep from doing it, with only modest success.

Regarding the selection of clubs in the U. S. Open Golf Tournament, not only do the players select the clubs with what they perceive to be the proper loft, but also, they have in mind the kind of shot they are going to play. They "see" the flight of the ball, where, and how it will land. The winner, Corey Pavin, is well-known as a "shot maker." That is, he can hit a ball high, low, with draw, or with fade. It is not surprising that he (or a player with these talents) won this tournament. Also, his puts were firm and assured.

Golf balls may have no preference in how they fly, but the makers of them do. It is possible to hit balls that will "tend" to take a high trajectory, a low trajectory, or medium. Also, balls are being made that will minimize hooks and slices. Alas, the players still must select clubs, imagine shots, execute them, chase after the ball, and hit it again.

I may take some time to think more of PCT and golf - thus combining two of my favorite things.

Later, Dan miller
Date: Mon, 19 Jun 1995 17:25:14 -0600

Subject: The Domains of reinforcement theory and PCT

[From Bill Powers (950619.1455 MDT)]

Some thoughts on reinforcement theory that must have passed through the minds of those who developed it, if not exactly in the order presented below. And following, a possible resolution of the conflict between reinforcement theory and control theory.

Consider this simple closed-loop representation of the reinforcement model:

```
              --------
  ------>|  ORG   |---->--
 |        --------        |
 |                        |
 |                        |
 |        --------        |
 R<----- |  ENV   |<------B
          --------
```

The organism outputs a behavior B with a reinforcing consequence R produced through an environmental link. The reinforcing consequence R enters the organism via senses, and the result is change in the behavior B. What kind of change in B follows from the reinforcement R?

If we think of B and R as events, then there can be no change in either B or R taken as individual occurrences. They either occur or they do not occur; there is no other choice. A lever is either pressed or it is not pressed (B); a reinforcer is either given as a consequence or it is not given (R). This would seem to leave reinforcements with no way of having any effect on behaviors. This is not what we want as a model of learning.

Suppose we thought of B and R as having variable magnitudes. Now we can have more or less of B, and more or less of R. This allows R to have an effect on B in terms of magnitude. If we assumed that an increase in R causes an increase in B and that R is proportional to B, then we would have

R1 = k1*B1
B2 = B1 + k2*R1.
R2 = k1*B2
B3 = B2 + k2*R2,

and so on. From this we would quickly conclude that the behavior B will either get larger and larger without bounds (k1*k2 > 1) or smaller and smaller until it is zero (k1*k2 < 1).

This model has no stable states other than infinity or zero. This is not what we want as a model of learning, either.

What we can do is go back to the idea that B and R are events that occur or don't occur, but consider the _probability_ that B will occur. Now we can say

B --> R (the behavior, if it occurs, leads to R)

pr{B} = pr{B} + k*R {where R is either 1 or 0)

This arrangement does not cause behavior to run away to infinity, because now the greatest possible probability is 1. If the behavior starts with a zero probability of occurring and the probability is increased every time R occurs, eventually the probability of B occurring will become 1, and B will occur every time R occurs.

This, then, gives us a basic model of learning with the right properties. At first there is no behavior or very little. When the behavior does appear, it produces a reinforcement, which in turn increases the probability that the same behavior will occur again. With enough repetition, the behavior will eventually occur every time the reinforcer occurs. Since the behavior always results in a reinforcement, the behavior will persist.

This model can now be expanded to include a discriminative stimulus S, which signals the conditions under which B will produce R. Now the probability that is altered by R is Pr{B|S}: the probability that occupance of the discriminative stimulus S will lead to behavior B. Again there is no runaway condition, because the maximum possible probability of any event is 1. The most that can happen even with continued reinforcement is that every time S occurs, B will occur.

If the probability in question is considered to be a probability _density_ (the probability of occurrence per unit time), the probability of B can be interpreted as the average interval between occurrences of B. If the probability density is high, B will occur after only a few time-increments; if low, B will occur only after many time increments. This leads to seeing the natural measure of behavior as _rate of repetition_, where the average rate of repetition is simply the reciprocal of the probability density.

Thus even though behaviors and reinforcements themselves are considered as unitary events which either occur or don't occur, we can find a continuous variable in terms of which to measure behavior: its rate of occurrence, which is closely related to the basic measure of probability of occurrence per unit time. The rate measure of behavior and reinforcement is thus the natural result of considering behaviors and reinforcements to be events, and of proposing that the effect of a reinforcement on the production of a behavior (or on the response to a discriminative stimulus) occurs via an effect on the probability-density of the responses. The use of probabilities is dictated by the fact that if the magnitude of reinforcement had an effect on the magnitude of behavior, the resulting model would have stable states only at zero and infinity.

A ratio schedule is one in which the occurrence of a reinforcer depends probabilistically on the occurrence of a behavior. If the probability is less than 1, we have a variable ratio schedule which requires a number of behavioral acts to produce one reinforcement. A fixed ratio schedule is a regular approximation to the variable-ratio schedule. Once again, there is a natural limit that prevents runaway: there can be no more than one reward per behavior; the probability of a reward given a behavior can be no greater than 1.

From here the development can be taken in many directions which will not be considered here.

We can see that as originally conceived, the model is a positive feedback model. The more reinforcement there is the more behavior there is, the more behavior there is the more reinforcement there is. The converse also holds true; if there is any lessening of either behavior or reinforcement, both variables must decrease with the only limit being zero.

To avoid this obviously inappropriate result, the meaning of "more" had to be modified so that the outcome was not a runaway condition. This is the function of the concept of increasing _probabilities_ rather than _magnitudes_. The model remains a positive feedback model, but the insertion of a probability puts a limit on the runaway condition where the probability becomes 1.

Let's go back to the original model, but this time add a reference signal and change the system to a negative feedback model:

```
                   R'*
                    |+
                --------
     ------>|  ORG   |---->--
    |           --------       |
    |                          |
    |                          |
    |           --------       |
    R<----- |  ENV   |<------B
                --------
```

The symbol R' (R-prime) now stands for a desired amount of R. While it is not evident in the diagram, we have also changed the sign of the effect of R on B. Now B is determined by the difference between R' and R: that is,

$$B = k1*(R' - R).$$

As a result, an increase in R will cause a decrease in B. As before, an increase in B will cause an increase in R:

$$R = k2*B$$

If we now consider that R and B are variable in magnitude, rather than being events, we can solve the above two equations for the steady-state result:

$$R = k2*(k1*(R' - R), \text{ or}$$

$$R = \frac{k1*k2}{1 + k1*k2} R'$$

If k1*k2 becomes much larger than one, we find that the result is

$$R = R'.$$

Now thinking of R and B in terms of magnitudes leads to a stable system. R will come to the magnitude specified by R', and B will be R'/k2. The only requirement is that k1*k2 be so large that we can neglect the 1 in the denominator.

Needless to say, realizing either of the above two models in a physical system requires adding some details: in the reinforcement model we must find a physical way of doing the equivalent of changing a probability. In the negative feedback model we must insert a filter that allows k1*k2 to be large without resulting in oscillations.

Under what conditions would we expect the reinforcement model to be appropriate? The answer can be found by asking under what conditions we could legitimately consider only the _occurrence_ of a behavior and a reinforcement without also considering their magnitudes.

When B is considered as an event, we have only the choice between that behavior being observed and its not being observed. This is equivalent to seeing either the behavior B that results in R, or SOME OTHER BEHAVIOR that does not result in the appearance of R. When an operant behavior is being acquired, this is exactly the situation that is seen. If the rat is not pressing the bar, it is doing something else. So to speak of the probability of the behavior is to speak of the probability that the _right_ behavior, among all those kinds of behavior that are possible, has occurred.

To say that a reinforcement increases the probability of a behavior thus can legitimately be taken as meaning that R increases the probability that the _right kind_ of behavior will appear. Instead of rearing up against the wall of the cage in one place, the rat rears up in another place such that the act of rearing up causes the lever to be depressed by the front paws. What we observe is an increase in the probability that the rat will rear up in that one critical position rather than in any other. Equivalently, since we can't directly observe probabilities, we can say that the rat's behavior changes so it shows an ever-greater proportion of rearing-up actions in the right place relative to all other places.

This is the kind of positive feedback situation in which there is a natural limit to the runaway condition: the animal can't spend more than all of its time performing the right kind of behavior to produce reinforcements. The reinforcement model is therefore applicable, if anywhere, to the process of _acquiring the right kind of behavior_ as opposed to behaviors that have no effect on producing reinforcements.

Once the right kind of behavior has been established, the reinforcement model has gone as far as it can go. It cannot also account for the fact that the animal comes to produce exactly the amount of behavior that is needed to bring the amount of reinforcement to the _right amount_. That process requires a negative feedback model with a reference signal, with provision for adjusting the feedback parameters for best control.

The one area of confusion that is left concerns measuring behavior in terms of rate of occurrence. This confusion shows up because in the attempt to explain performance as well as acquisition of behavior, experimenter-theoreticians set up experiments in which it was impossible, even when the right behavior had been acquired, to vary the _amount_ of effect produced by a single behavior. The use of levers and keys effectively removed any ability of an animal to have a quantitative effect on the reinforcement by varying its efforts.

Even in this artificial situation, an animal could vary its behavior as a way of varying the _average_ amount of obtained reinforcement. Once the right behavior had been found, pressing a key or lever, the animal (or at least some animals) could now vary the amount of received reinforcement by varying the rate at which it pressed the key or lever.

So a negative feedback model can be set up in which R and B are both measured in terms of rates, with R having an average perceived effect depending on the rate of delivery of reinforcements and the internal rate of decay of effects of reinforcements. This internal perceptual effect could then be compared with the desired effect, and the difference could be converted proportionally into a rate of generation of bar or key pressing acts.

The reinforcement model can _also_ be set up to produce a continuous relationship between rate of bar pressing and rate of reinforcement, through the intermediate step of effects of reinforcement on probability of behavior per unit time. However, this creates a conflict between the reinforcement model and the negative feedback model. The reinforcement model says that an increase in reinforcement rate must produce an increase in the probability (frequency) of behavior, while the negative feedback model says that an increase in reinforcement rate must produce a _decrease_ in the behavior rate. If this conflict can be resolved, there will be no further difficulties between the reinforcement model and the negative feedback control system model. This is not to say that neither model will be further modified on the basis of other considerations, but at least this direct contradiction will have been removed.

The contradiction can be removed by saying that the reinforcement model applies strictly to the process of increasing the relative frequency with which the right behavior occurs, in comparison with all non-reinforcement-producing behaviors that might also occur. Once the right behavior has been acquired, the negative feedback control system model applies to the process of creating the desired amount of reinforcement.

In the data that appear ambiguous, supporting the reinforcement model at one extreme and the control model at the other, the difference can now be explained easily, and in a way that can be tested against experimental data. Where reinforcement rates are low enough, the reinforcement model applies, and the animal begins to search for other behaviors that will more reliably produce the reinforcer. This means that other behaviors beside the bar-pressing will be seen, and proportionally less time will be spent doing the right behavior. This shows up as an apparent drop-off of behavior with decreasing reinforcement rates, or an apparent increase in behavior rate (of the kind being measured) with increasing reinforcement rates. Where reinforcement rates are high enough, the animal essentially always uses the right behavior, and controls the amount of received reinforcer near a specific reference level by varying its rate of behavior: now an increase in reinforcement rate goes with a decrease in behavior rate.

Best to all EABers and PCTers, Bill P.

Date: Mon, 19 Jun 1995 23:42:51 -0400

Subject: Then a miracle occurs [FROM: Dennis Delprato (950619)]

>Bill Powers (950618.1430 MDT)

> What has me climbing the wall is the use of the term "selection by consequences" when to me it is obvious that nothing of the sort is going on.

> I've been watching the U.S. Open golf tournament on TV. Shinnecock is a tough course and club selection for short approach shots is critical. If the club is too lofted the wind makes the ball drift, and if not lofted enough the ball runs when it hits the hard green as well as falling short. So what selects the club? The consequence of drifting in the wind or falling short, or THE PLAYER? I contend that it is the player who does the selecting, using information about previous experience to be sure, but still using a process of selection that lies in the player and not in the ball.

> The behavior of a golf ball is just not the sort of thing that can "select." Selection of behavior is a process of weighing many factors and choosing an action that is estimated -- by a system capable of making estimates -- to produce a desired result. The ball has no preference for how it will behave and mechanism for making estimates or choosing on that basis; only the player can have such a preference, or draw conclusions from present-time data.

There appears to be a misunderstanding here; in reply, I'll speak "operantly" to the best of my ability and with brevity in view: It seems like you are interested in understanding and explaining a golfer's behavior of asking for and using a particular club on a particular occasion. Young fellow, I assure you that the answer is not a cognitive expectancy, a schemata, any sort of underlying motive, a strong S-R bond, or any other purely hypothetical entity that your elders may have taught you. The behavior also is not a manifestation of an inherited trait. The answer is in the golfer's history of reinforcement. Consider a particular set of conditions [wind conditions, lie (isn't this a golf descriptor?), distance from hole, perhaps organic condition of golfer (e.g., degree of fatigue), condition of green, slope of green, ...]. In the past, when the golfer has used a particular club, they were more reinforced (positive reinforcement: ball went closer to hole; negative reinforcement: ball did not end up in bad place like 'the rough) than if they used other clubs. In slightly more complete terms, we might say the golfer's responses were 'differentially reinforced': Response-1 -->positive reinforcers presented, negative reinforcers not presented; Responses other than 1--> positive reinforcers less likely, negative reinforcers more likely. (--> means followed by where the right side of --> indicates what we call 'consequences (of responding). The basic principle being obeyed here is that responses followed by consequences that are reinforcers (presentation of positive reinforcers, termination or withholding of negative reinforcers) are selected. That is, they are more likely to occur in the future under the same or similar conditions (ess Ds, discriminative stimuli).

Thus, we see that you are basically correct: THE PLAYER selects. But why does THE PLAYER select? Because of THE PLAYER'S history of interactions with their environment. This history changes THE PLAYER in ways that physiology ultimately will describe. All we know now is that reinforcement histories change organisms organically and in this way behavior is modified. It appears that from my operant perspective I find the environmental consequences of past responses, along with current environmental (including organic ones) to be of particular interest. You seem to looking more than I into the question of what is happening organically in conjunction with all this. Perhaps we could get together some time.

>> One place to look might very well be the work on feedback functions. It seems to me that the molar behaviorists (where one finds feedback functions) have departed from Skinnerian "molecular" selection by consequences. They say, I think, "When responses are modified by contiguous relationships between responses and consequences, it is not what one sees directly (contiguous relationships) but what one does not see directly that is more generally important for describing what is going on."

> Left out of this way of putting it is the actual mechanism that does the modifying. That mechanism has to lie within the organism -- where else could it possibly be? The mere fact that a response has a certain consequence can't modify anything else. It's just a fact, something that happened. I don't think that proposing that contiguous relationships can modify responses is any better than saying that consequences can select behavior. Neither one makes any sense to me. How on earth could a contiguous relationship do anything but be a contiguous relationship? Wherein lies its ability to modify a response?

> Help me out, Dennis. Am I the only one in the world who sees statements like this as invoking magic? Here is a consequence; here is a contiguous relationship. Then a miracle occurs, and responses are modified or behavior is selected. Doesn't anyone else see something vital missing from these statements?

Is the operant view of the locus of the of the 'actual mechanism' of modifying now clearer? They focus on reinforcement history, i.e., actual response --> consequence relationships in life of individual. Skinner agrees that response --> consequence histories change organisms organically. He holds that despite this, the study of organism-environment interactions is important in its own right and if done well will provide fundamental material for those tracing what is happening to the organism at the organic level as a function of particular environmental histories.

On the miracle: I believe Skinner might ask if we believe Darwin was invoking miracles with the principle of selection by consequences (natural selection). K. U. and Tom Smith say neoDarwinians are, but that is another (cybernetic) story. To Skinner, natural selection is a basic law (causal mode, even) of biological evolution, the ontogeny of individual behavior, and cultural practices. We are not dealing with anything complicated with 'selection by consequences.' The alternative is taken to be "the causality of classical mechanics" or "'selection pressure,' which appears to convert selection into something that forces change. ... [with a more serious example being] the metaphor of storage" (Skinner, 'Selection by Consequences, _Science_, 1981, v. 213, p. 503). Basically, selection by consequences seems to be more simplistic or incomplete, rather than mystical -- what we find now is found because there was something functional or adaptive about it and that is why we find it.

Bill, you didn't comment on my suggestion that PCT modelers base models on PCT units not on units from non-PCT areas such as operant psychology. Perhaps when Bruce Abbott returns he will be able to determine if I am making any sense at all, because he is well-schooled at using the operant fundamental units (depending on one's view: respondent, operant, discriminated operant). In other words, one not practiced at using the operant units in technical ways could be enticed to use them unknowingly in models and thus develop nothing more than another operant model.

Dennis Delprato

Date: Mon, 19 Jun 1995 23:44:32 -0400
Subject: Re: The Domains of reinforcement theory and PCT

<[Bill Leach 950619.23:09 U.S. Eastern Time Zone]
>[From Bill Powers (950619.1455 MDT)]

An excellent posting I think... though I suppose that we will really have to wait for Bruce to return and see how he perceives it.

Something that occurred to me as I was trying to "think like the devil's advocate" upon reading your posting:

Regardless of what "errors" may have existed in the thinking associated with the "control" test that Bruce described, I believe that an important concept might have been missed (or at least the concept's real significance might have been -- and assuredly was by me anyway).

It seems that the discussion centered upon the test methodology and of course the related testing performed in an attempt to confirm assumptions.

Though the purpose of the test was talked about (and I believe that Rick even mentioned "search for controlled variable"), it seems that it was not really noticed that "trying to determine of the rats preferred the ability to control the shock conditions" is clearly NOT a search for a re-enforcer NOR is it a search for a particular ACTION or OUTPUT (behavior in the EAB sense, I believe).

In other words, the explicit function of the test was to determine if a rat had a particular PURPOSE (i.e.: PCT reference implying a controlled variable). It seems that this is the case, even if the researchers were not themselves aware of the vast difference between trying to determine what the rat is observed to do as opposed to WHY (from the rat's own perspective) it is doing anything.

The other related tests were specifically intended to insure that observed behavior would not be do to other possible goals of the rat. And again, regardless of the possibility of serious "flaws" in both the specific methodology and thinking with regard to experimental design, the basic idea was right whether it was recognized to be so for the right reasons or not.

Once again, I might be engaging in "wishful" thinking but if the above is even close then EAB is at least heading in the right direction for the recognition of the validity of PCT.

-bill

Date: Tue, 20 Jun 1995 00:15:57 -0400
Subject: Re: Then a miracle occurs

<[Bill Leach 950619.23:53 U.S. Eastern Time Zone]
>[FROM: Dennis Delprato (950619)]

> Perhaps when Bruce Abbott returns he will be able to determine if I am making any sense at all, because he is well-schooled ...

Yes SIR! I for one will be very interested in what Bruce has to say about this particular posting.

The "wording" might "tend to drive a PCTer up the wall" but the meaning of your golf example sure sounds like "learned control" to me.

Your "reinforcement histories"; "more" "positive reinforcers" and "less" "negative reinforcers" (hell of a pair of terms there, those two!) is a case of "control attempt with little or no error" and; "less" "positive reinforcers" and "more" "negative reinforcers" is a case of "control attempt with high error possibly even control failure" (ta da! Reorganization to the rescue?)

One place where there has not been a great deal of discussion with respect to "reinforcement" is in terms of memory. There _seems_ to be some correlation between repetitive experience and effectiveness of memory AT TIMES (at other times, it appears that a single experience that "was long forgotten" will in fact remain available.

Similarly (as a person that has recently just "half-heartedly" taken up golf), many "mechanical skills" seem to require repeated performance and "tuning" to eventually be performed well (yet again, occasionally there is that example where "first try" is correct).

If this is the sort of thing that "reinforcement" actually refers to; that is, the "consequences of a _control action attempt upon the controlled variable_ AS PERCEIVED by the organism (i: control error does or does not exist) AND NOT just the consequences as observed by others then maybe EAB is not as greatly different as it appears.

-bill

Date: Mon, 19 Jun 1995 21:56:49 -0700
Subject: R. coli II: With a vengeance

[From Rick Marken (950619.2200)]

I've been wrestling with the angel of reinforcement theory all weekend. I think I now know how to deal with reinforcement theory on its own terms. That is, I think I can show why a reinforcement model, in which response probabilities are changed by the consequences of responses, cannot be a model of learning or purposeful behavior. The reason is simple: consequences really DO select behavior in reinforcement theory and consequences just don't care what the organism does.

I have developed a new, even simpler version of R. coli to illustrate this point. In this experiment there is a cursor and a target. The position of the cursor is x; the position of the target is t. The cursor moves at a constant rate in one direction until the subject presses the mouse button. After a press, there is a 50-50 chance that the cursor reverses direction. The subject's task is to keep the cursor near the target and this is easily done.

The reinforcement model of this task is simple. There are two discriminative stimuli; S+ is cursor movement toward the target; S- is cursor movement away from the target. There are three consequences of a response: R+ is a reinforcement and it occurs only when the cursor changes from moving away to moving toward the target (so a reinforcement can only occur when the cursor is moving away from the target): R- is a punishment and it occurs only when the cursor changes from moving toward to moving away from the target (so punishment can only occur when the cursor is moving toward the target): R0 is non-reinforcement and it occurs only when a response results in no change in the direction of the cursor; so a non-reinforcement happens after half the responses when the cursor is moving toward the target and after half the responses when the cursor is moving away

This reinforcement model is the one dimensional equivalent of Bruce Abbott's version of the reinforcement model of E. coli. This model will keep the cursor near the target if the probability of responding when the cursor is{moving away from the target, P(R|S-), is close to 1.0 and the probability of responding when the cursor is moving toward the target, P(R|S+), is close to zero. Optimal control occurs when P(R|S-) = 1.0 and P(R|S+) = 0.0.

Now, what was obvious to Bruce Abbott (950616.0955 EST) about my previous version of R. coli is still obvious about this version:

> Responses are emitted until the cursor moves to target or the behavior extinguishes from lack of reinforcement (whichever comes first). Occasional success reinforces button-pressing, thus maintaining the behavior.

This verbal description of the situation implies that P(R|S-) is incremented when R+ occurs after a response, R, in the presence of S-. That is:

$$P(R|S-)_{n+1} = P(R|S-)_n + \text{alpha}\ (1-P(R|S-)) \qquad (1)$$

where alpha is the learning rate parameter and n is the trial number. If this were all there were to the model, everything would work fine; P(R|S-) would go to 1.0 after a certain number of trials and responses would control the cursor.

It wasn't until Saturday night that I realized that the model in equation (1) can't be the whole story; the result of equation (1) is a model that keeps pressing, with a probability of 1.0, when S- is present; the pressing never extinguishes, even if you change the situation so that the cursor moves away from the target after every press (S- is always present).

This is obviously NOT the way the model should work -- as Bruce says, behavior should extinguish when there is lack of reinforcement. If responses in the presence of S- continuously produce R0 (non-reinforcement) as a consequence and the model keeps responding to S- then behavior (responding) is clearly no longer being selected by its consequences!!

Reinforcement theory, therefore, implies that P(R|S-) must decrease when R0 occurs after a response, R, in the presence of S-. The question is how this decrease should be implemented. I thought the following seemed reasonable:

$$P(R|S-)_{n+1} = P(R|S-)_{n} - \text{beta } (P(R|S-)) \qquad (2)$$

Equation (2) applies when a response in the presence of S- is followed by no reinforcement, R0. P(R|S-) decreases from its current value at the rate beta. Both equations (1) and (2) are required if the reinforcement model is to be a selection by consequences model; one that is consistent with Bruce Abbott's verbal description of a reinforcement model:

> Responses are emitted until the cursor moves to target or the behavior extinguishes from lack of reinforcement (whichever comes first). Occasional success reinforces button-pressing, thus maintaining the behavior.

The parameters of the reinforcement model, alpha and beta, have a lot to do with how the model works. If alpha >> beta the model learns to control quickly but takes MANY trials to extinguish. If beta >= alpha the model never learns to control.

The sensitivity of the reinforcement model's behavior to alpha and beta makes sense because these parameters determine how the environment (the consequences R+ and R0) influences behavior. If alpha >> beta then R+ has a big effect on responding; this produces a system that persists in responding to S- even when such responding no longer produces what an outside observer can see as the appropriate results. If beta >> alpha the system never responds in a way that an observer would see as appropriate.

So the parameters of the reinforcement model determine only how it is "pushed around" by the environment -- as one would expect from a model whose behavior is controlled by consequences. Certain kinds of pushing around (like the kind that occurs when alpha >> beta) can look very useful because the system is pushed into responding like a control system (a control system that cannot change its reference for the controlled variable); but this control system type behavior only happens when the system is designed to let the environment push it around in a certain way. When alpha = beta the system never gets pushed into being a control system; it just responds randomly -- and could care less about the fact that it lets the cursor wander aimlessly up and down the computer screen.

An outsider observer can adjust the parameters of a reinforcement model so that it is pushed into responding in a particular way in a particular environment. But, in general, the reinforcement model doesn't really learn to control in different environments; it just goes with the flow. In some cases, the flow makes it respond in such a way that it controls a variable, and sometimes the flow takes away control.

Best Rick

Date: Tue, 20 Jun 1995 11:03:32 -0600
Subject: Operant psychology

[From Bill Powers (950620.1235 MDT)]

Dennis Delprato (950619) --

> There appears to be a misunderstanding here; in reply, I'll speak "operantly" to the best of my ability and with brevity in view: It seems like you are interested in understanding and explaining a golfer's behavior of asking for and using a particular club on a particular occasion. Young fellow ...

Well, I might forgive some of what follows "young fellow," but not what comes before that reinforcer (or much that comes after it).

No, I'm not interested in understanding the golfer's behavior of asking for and using a particular club. What I'm interested in is what operations go on between the "history of reinforcement" and the golfer's reaching for his eight-iron. That is where reinforcement theory relies on magic.

> Young fellow, I assure you that the answer is not a cognitive expectancy, a schemata, any sort of underlying motive, a strong S-R bond, or any other purely hypothetical entity that your elders may have taught you. The behavior also is not a manifestation of an inherited trait.

That's a strange thing to say to me. What about hypothetical entities that I made up myself, such as input functions, comparators, and output functions? Or are you playing Devil's Advocate?

> In the past, when the golfer has used a particular club, they were more reinforced (positive reinforcement: ball went closer to hole; negative reinforcement: ball did not end up in bad place like 'the rough) than if they used other clubs.

I would rather see the same events described more neutrally, without asserting that there was anything "reinforcing" about them one way or the other. I believe it's important to state observations without overlaying them with unprovable theoretical assumptions, such as the assumption that some sort of unproven "reinforcing" effect occurred. If you assume your theory as part of the statement of what you observe, the reports are biased against alternative interpretations.

> The basic principle being obeyed here is that responses followed by consequences that are reinforcers (presentation of positive reinforcers, termination or withholding of negative reinforcers) are selected.

That is the claim, I agree.

> That is, they are more likely to occur in the future under the same or similar conditions (ess Ds, discriminative stimuli).

Could it not be that the consequences are more likely to repeat (in the absence of disturbances) when the behavior creating them repeats? Where is there any evidence that it is the reinforcer that is causing the behavior to repeat? We can see exactly how the changed behavior produces the changed consequences by physical means. We cannot _observe_ any effect of the consequences on the behavior.

And who says that repeating consequences result from repeating behaviors? That is a myth. Repeating consequences, in a natural environment, result from _varying_ behaviors.

> Thus, we see that you are basically correct: THE PLAYER selects. But why does THE PLAYER select? Because of THE PLAYER'S history of interactions with the environment. This history changes THE PLAYER in ways that physiology ultimately will describe.

There is no doubt that a history existed; what is in doubt is whether this history has any kind of controlling effect on the selection. You are committing the _post hoc, ergo propter hoc_ fallacy: after which, therefore because of which. It is possible that the history reflects a continuing process of adjusting behavior to make the consequences come to a preselected condition, in which case the history is simply a side- effect of the process.

> You seem to looking more than I into the question of what is happening organically in conjunction with all this. Perhaps we could get together some time.

You don't seem to be looking at all into what is happening organically. The PCT model is a way of trying to guess at what is happening organically, the guesswork being supported to some extent by physiological observations. What we find in the PCT model is a physical way of incorporating intentions and purposes into behavior that leaves the environment with essentially no control over behavior. Even though the physiological evidence in support of this view is far from complete, it at least exists, which is more than can be said for the view you describe.

> Is the operant view of the locus of the of the 'actual mechanism' of modifying now clearer? They focus on reinforcement history, i.e., actual response-->consequence relationships in life of individual. Skinner agrees that response-->consequence histories change organisms organically. He hold that despite this, the study of organism-environment interactions is important in its own right and if done well will provide fundamental material for those tracing what is happening to the organism at the organic level as a function of particular environmental histories.

Of course the study of organism-environment interactions is important in its own right. Every explanation of behavior has to have some kind of observations to explain. The better one observes, the more important the test of the theory becomes.

But your restatement of the observations does not make the locus of the actual mechanism any clearer. You are only describing what observed variables do, not the physical system that lies between them and imposes the relationships we see -- i.e., the mechanism itself. And you also mingle the straight observations with subtle assertions about invisible causal links.

To say that histories "change organisms organically" is to say precisely nothing, because nothing is said about WHAT organic changes, with what significance, occur. It's like explaining the action of a drug on the body by saying that it's "chemical."

> To Skinner, natural selection is a basic law (causal mode, even) of biological evolution, the ontogeny of individual behavior, and cultural practices.

Natural selection is not a basic law, it's a way of characterizing in a couple of words a whole interaction between a species and its environment. The use of the word "selection" is completely metaphorical; taken literally, it's an effect without a cause (that is, selection occurs, but there is no selector).

> Bill, you didn't comment on my suggestion that PCT modelers base models on PCT units not on units from non-PCT areas such as operant psychology.

We would be happy to do so, if operant psychologists didn't keep insisting that they have already explained the same phenomena we explain, and if their explanations of phenomena we haven't yet explained were credible.

Dennis, all you have shown in your reply is that Skinner's interpretation is both self-consistent and compelling to some. But it relies on continual assertion of unobservable relationships among variables. These assertions are woven into descriptions in such a subtle way that the fact that an unjustified assertion is being made is almost invisible -- particularly to those who have bought into the system. And any alternative explanations are automatically dismissed as unthinkable or unscientific, so it is never necessary to show what is actually wrong with them.

Consider this description:

Every time the player uses an eight-iron under windless conditions and with a full swing, the ball travels about 160 yards (don't know if that number is right but it's close). When the ball is about 160 yards from the green and the same conditions hold, the player often uses an eight-iron. This situation evolved from one in which the player used a different iron under the same conditions, with different results. As the club-selection changed, the results changed. Eventually the club-selection came to be what we observe, most of the time, and the results came to be, most of the time, those that follow from the club selection.

Then add this:

The reason that the player came to use the eight-iron under those conditions is that the resulting position of the ball in combination with wind conditions has had an effect in the past on the player's nervous system that has caused the nervous system to increase the probability of selecting an eight-iron.

Now it's clear that the plain description of historical events is quite separate from the claim that the consequence of the behavior is having an effect that causes that behavior to become more likely. There is, in fact, no observational support for this conjecture about why the player selects an eight-iron. The reinforcement hypothesis is in the same class as hypothetical motives, expectancies, and so on.

Of course if such a reinforcing effect did occur, it would nicely explain why the player uses the eight-iron. If a "trait" did exist, it would nicely explain why a person shows certain characteristic behaviors. But that is not how scientific explanations work. In order to justify that explanation, you have to look at something other than the result it explains. You need _independent confirmation_ that such an effect exists. Otherwise, you have a plausible explanation but not a scientific explanation. You have an untested hypothesis. That hypothesis could be right, but it could also be wrong. Nobody will know until it's tested.

How would you explain all of your lights suddenly going out? You might say, "There's been a general power failure." True enough, if a general power failure had occurred, all your lights would go out. But that does not show that your explanation is correct. You need some kind of independent confirmation of the hypothesis, other than the fact that your lights are out. You might, for example, go outside and see if all your neighbors' lights are out, too. If you can see only one or two neighbors' houses, you might want to go further and find out if they are actually at home, or even check their power meters to see if the rotor is turning. In other words, the original hypothesis can't be checked by seeing if its logical consequence has occurred, the very consequence you're trying to explain. The credibility of your explanations depends on how much checking for independent confirmation you do.

If you go outside and see that your neighbors' houses are all lit at least with night-lights, your hypothesis collapses immediately and you have to seek an alternative explanation, like a major short-circuit in your house wiring.

Skinner came up with a big complex explanation of behavior that is consistent with what we see (up to a point, but that's another subject). If his explanation is correct, it would account for many observations. But his whole system is an untested hypothesis that requires independent confirmation before it can be accepted as a scientific explanation. The whole explanation could be wrong.

Among proponents of Skinner's interpretations, I see nothing in the way of seeking independent confirmation. Instead what I see is a continual effort to show that the proffered explanations would indeed predict what we observe. If some difficulty arises, like data that don't seem to jibe with the basic picture, the reaction is simply to start manipulating the basic variables of reinforcement theory until a combination is found that will lead to at least a qualitative approximation of what is observed. Since in most situations many variables can be identified, showing many different relationships to behavior, one can always find a variable to play the role of the reinforcer or the discriminative stimulus, and that is usually sufficient to construct a plausible explanation -- one that fits what was observed.

What I don't see is any attempt to challenge the theory -- to ask how we could verify by some means other than pointing to the phenomenon to be explained that a reinforcing effect occurred, or that a discriminative stimulus was in fact detected by the nervous system of the organism. The general attitude I find is that no other interpretation of the observations is even conceivable within a scientific framework.

Best, Bill P.

Date: Tue, 20 Jun 1995 10:28:49 -0700
Subject: Reinforcement Theory: As You Like It

[From Rick Marken (950620.1030) --

I think it is significant that reinforcement theorists must develop a new version of their theory each time a new behavioral situation is described. For example, in a simple operant conditioning experiment, the occurrence of food following a bar press is considered a reinforcement. I have never heard reinforcement described as the change from "no food" to "food". So, when I developed the E. coli demo it seemed reasonable to assume that the direction of movement following a press was the reinforcement; like the food in the basic operant conditioning experiment, direction of movement is a direct result of a response. But, in order to make the reinforcement model of E. coli behavior work, it was necessary to redefine reinforcement as _change_ in the direction of movement after a press; change toward the target was the reinforcement.

It is also sometimes necessary to change what was considered a non-reinforcer or punisher in certain experiments. This is what happened in the first version of my R. coli demo. In that demo, a press leading to a change from non-target to target is considered a reinforcer but a change from one non-target to another could not be considered a punishment because, if it were, the reinforcement model wouldn't work.

The events that are considered reinforcement, punishments, etc must be changed to account for behaviors in different situations because the behavior of the reinforcement model is determined by the situation. The model will only work (account for the actual behavior observed) if the modeler can pick out aspects of the environment that will produce the right behavior. This, of course, means that the reinforcement model cannot fail as long as one can find appropriate environmental inputs (appropriate consequences of responding) for the model.

Martin Taylor (950620 11:10) has noticed this characteristic of reinforcement theory. He asks:

> wouldn't maintenance of the direction of the cursor away from the target be negative reinforcement, and maintenance of its direction toward the target be positive reinforcement?

(I presume Martin is referring to the _occurrence_ (not maintenance) of a particular direction of cursor movement). I understand Martin's puzzlement. But I have gotten used to the idea that reinforcement theorists identify events as reinforcers, punishers, etc. as necessary to produce the intended results. The direction of movement of the cursor after a press is considered a positive or negative reinforcer, a punishment or a non-reinforcer depending on whether considering it so contributes to the appropriate behavior of the model.

So, in answer to Gary Cziko (950614.1920 GMT) who said:

> Your original E. coli demo may not have done what you wanted it to do, but I think your approach was a good one--finding the simplest example of goal-oriented behavior that could not be accounted for by reinforcement theory.

I can now see that there is probably no way to show that goal-oriented behavior cannot be accounted for by reinforcement theory. As long as you can come up with some aspect of the environment that covaries along with the behavior being produced you can let that aspect of the environment increase of decrease the probability of response, as necessary, to get the result you want.

This is the "power of reinforcement theory" to which Bill Powers (950617.IHateToSay) described. But I think reinforcement theory only has this power (to account for any behavior) as long as behavior is dealt with in terms of probabilities of response. Obviously, reinforcement theory could not account for the quantitative details of the typical compensatory tracking experiment.

Of course, the power of reinforcement theory exposes its basic flaw. The power of reinforcement theory is achieved by looking for aspects of the environment that can account for the behavior we see. But this means that the same aspect of the environment must be treated as a reinforcer in some circumstances and a non-reinforcer or punisher in others. Since the theory provide no independent means of measuring the reinforcingness or punishingness of the environment, we are free to ascribe these changing characteristics to the environment as necessary to explain behavior. This is not science; it is tautology.

Best Rick

Date: Tue, 20 Jun 1995 12:03:48 -0700
Subject: Kibbitzing

[From Rick Marken (950620.1200)]

Bill Powers (950618.1430 MDT) --

> So what selects the club? The consequence of drifting in the wind or falling short, or THE PLAYER? I contend that it is the player who does the selecting ...

Hans Blom (950620)--

> Watch out, this is a rat's nest... Discussions about "causes" and "effects" can be found in philosophy, but not with well-defined results.

Um. Bill was talking about _selection_ Hans, not cause and effect.

Bill Powers (950618.1430 MDT) --

> The behavior of a golf ball is just not the sort of thing that can "select."

Hans Blom (950620)--

> Why not? If the player is perfect (his world-model has fully con- verged), and even if he is not, he is purely REACTIVE: the (perception of the) environment dictates what he is going to do.

A golfer is purely REACTIVE? Not in my world-model;-)

> In the case of a golf ball, you accept determinism. In organisms, you do not. I wonder why.

Wonder no more. The cause-effect relationships in a control system occur in a closed-loop.

Bill Powers (950620.1235 MDT)--

> Among proponents of Skinner's interpretations, I see nothing in the way of seeking independent confirmation. Instead what I see is a continual effort to show that the proffered explanations would indeed predict what we observe. If some difficulty arises, like data that don't seem to jibe with the basic picture, the reaction is simply to start manipulating the basic variables of reinforcement theory until a combination is found that will lead to at least a qualitative approximation of what is observed. Since in most situations many variables can be identified, showing many different relationships to behavior, one can always find a variable to play the role of the reinforcer or the discriminative stimulus, and that is usually sufficient to construct a plausible explanation -- one that fits what was observed.

> What I don't see is any attempt to challenge the theory -- to ask how we could verify by some means other than pointing to the phenomenon to be explained that a reinforcing effect occurred, or that a discriminative stimulus was in fact detected by the nervous system of the organism. The general attitude I find is that no other interpretation of the observations is even conceivable within a scientific framework.

Ah. Great minds (Rick Marken (950620.1030)) think alike;-)

Best Rick

Date: Tue, 20 Jun 1995 22:39:21 -0400
Subject: Operant psychology

[FROM: Dennis Delprato (950620)]

>Bill Powers (950620.1235 MDT)

>>Dennis Delprato (950619) --

>> There appears to be a misunderstanding here; in reply, I'll speak "operantly" to the best of my ability and with brevity in view: It seems like you are interested in understanding and explaining a golfer's behavior of asking for and using a particular club on a particular occasion. Young fellow ...

> Well, I might forgive some of what follows "young fellow," but not what comes before that reinforcer (or much that comes after it).

> No, I'm not interested in understanding the golfer's behavior of asking for and using a particular club. What I'm interested in is what operations go on between the "history of reinforcement" and the golfer's reaching for his eight-iron. That is where reinforcement theory relies on magic.

Your interests lie where Skinner acknowledged the ultimate answers are to be found: "The physiologist of the future will tell us what is happening inside the behaving organism. His account will be an important advance over a behavioral analysis, because the latter is necessarily 'historical'--that is to say, it is confined to functional relations showing temporal gaps. Something is done today which affects the behavior of an organism tomorrow. No matter how clearly that fact can be established, a step is missing, and we must wait for the physiologist of the future to supply it." (Skinner, _About Behaviorism_ (College ed.), 1974, p. 215)

Does not PCT eventually get to the physiological level? Why take PCT as a competitor to operant theory? Operant theory never was intended to 'get beneath the surface' of organism-environment interchanges. On the contrary, Skinner held that if all we could do was speculate about what was going on 'beneath the surface' then assume we were understanding something, we'd be best off to stay at the surface for a while, find order here, then we'd have some established findings to help guide the search for more complete understanding. Certainly, once research, not speculation, goes underneath the surface, we may find that certain 'knowledge' is not so certain after all. So be science. Again, in my view, not Skinner's (I don't know what his was), PCT and operant theory/research are not competitors. If PCT turns out to be more "right" than operant theory, it will because PCT takes us further into the details and intricacies of psychological events.

>> Young fellow, I assure you that the answer is not a cognitive expectancy, a schemata, any sort of underlying motive, a strong S-R bond, or any other purely hypothetical entity that your elders may have taught you. The behavior also is not a manifestation of an inherited trait.

> That's a strange thing to say to me. What about hypothetical entities that I made up myself, such as input functions, comparators, and output functions? Or are you playing Devil's Advocate?

Intent: to show that Skinner viewed operant theory as rejecting lineal causality and nonspatiotemporal explanatory constructs. All science requires constructs--PCT constructs are the kind needed in a science of behavior. They would not be the kind needed if they were handed down from cultural tradition, speculative, far-fetched analogies, and the like. It seems to me that PCT constructs are derived from expert observers interacting with things in the world-- this is how proper scientific constructs are derived.

>> In the past, when the golfer has used a particular club, they were more reinforced (positive reinforcement: ball went closer to hole; negative reinforcement: ball did not end up in bad place like 'the rough) than if they used other clubs.

> I would rather see the same events described more neutrally, without asserting that there was anything "reinforcing" about them one way or the other. I believe it's important to state observations without overlaying them with unprovable theoretical assumptions, such as the assumption that some sort of unproven "reinforcing" effect occurred. If you assume your theory as part of the statement of what you observe, the reports are biased against alternative interpretations.

When I trace the above to events and relations between events, I find the statement more of another way of describing what happened than asking us to make 'unprovable assumptions.' True, they then go on and talk as though the sort of description above explains everything--this is an assumption.

>> The basic principle being obeyed here is that responses followed by consequences that are reinforcers (presentation of positive reinforcers, termination or withholding of negative reinforcers) are selected.

> That is the claim, I agree.

>> That is, they are more likely to occur in the future under the same or similar conditions (ess Ds, discriminative stimuli).

> Could it not be that the consequences are more likely to repeat (in the absence of disturbances) when the behavior creating them repeats? Where is there any evidence that it is the reinforcer that is causing the behavior to repeat? We can see exactly how the changed behavior produces the changed consequences by physical means. We cannot _observe_ any effect of the consequences on the behavior.

They ask 'What are the critical variables that are responsible for the to-be-explained behavior?' Answer for operant behavior is that they are in the history of Response--> Consequences. Thus, the reinforcer itself does not cause responses to repeat. What can be observed, it is argued, is the organism's history of Response--> Consequence experiences; furthermore, these can be experimentally manipulated to test if such a history does contribute to responses/ behavior as 'dependent' variables. (Note for a complete even operant science, respondent interactions are recognized as well as operant ones.)

> And who says that repeating consequences result from repeating behaviors? That is a myth. Repeating consequences, in a natural environment, result from _varying_ behaviors.

Skinner addresses this with the distinction between response instance and response class. The former is particular and not repeated. The latter is functionally defined and comprised of response instances. In his terminology, 'varying behaviors' = 'response class.'

>> Thus, we see that you are basically correct: THE PLAYER selects. But why does THE PLAYER select? Because of THE PLAYER'S history of interactions with the environment. This history changes THE PLAYER in ways that physiology ultimately will describe.

> There is no doubt that a history existed; what is in doubt is whether this history has any kind of controlling effect on the selection. You are committing the _post hoc, ergo propter hoc_ fallacy: after which, therefore because of which. It is possible that the history reflects a continuing process of adjusting behavior to make the consequences come to a preselected condition, in which case the history is simply a side-effect of the process.

[The 'you' is my attempt to transmit operant theory] They use 'history' in a temporal sense--what happened previously in time. If you are saying that what happens to an organism historically is a function of the organism's behavior, operant theorists would agree.

>> You seem to looking more than I into the question of what is happening organically in conjunction with all this. Perhaps we could get together some time.

> You don't seem to be looking at all into what is happening organically. The PCT model is a way of trying to guess at what is happening organically, the guesswork being supported to some extent by physiological observations. What we find in the PCT model is a physical way of incorporating intentions and purposes into behavior that leaves the environment with essentially no control over behavior. Even though the physiological evidence in support of this view is far from complete, it at least exists, which is more than can be said for the view you describe.

Skinner agrees that he _is not_ looking into what is happening organically. See above. You have stated what I see as the gap that PCT can fill and why I think PCT theorists should quit worrying about reinforcement theory. When PCT is developed, it will explain all the enduring relationships discovered by operant psychology and more. Is there anything in operant psychology for the PCT researcher to be concerned with, then? Yes, are there any reproducible findings? Are there any operations/ procedures that yield predictable results? There are. Just be concerned with these. These are fodder for PCT, not what people _say_ about the relationships between their experimental operations and their findings. Who cares what the _say_?

>> Is the operant view of the locus of the of the 'actual mechanism' of modifying now clearer? They focus on reinforcement history, i.e., actual response--consequence relationships in life of individual. Skinner agrees that response--consequence histories change organisms organically. He hold that despite this, the study of organism-environment interactions is important in its own right and if done well will provide fundamental material for those tracing what is happening to the organism at the organic level as a function of particular environmental histories.

> Of course the study of organism-environment interactions is important in its own right. Every explanation of behavior has to have some kind of observations to explain. The better one observes, the more important the test of the theory becomes.

> But your restatement of the observations does not make the locus of the actual mechanism any clearer. You are only describing what observed variables do, not the physical system that lies between them and imposes the relationships we see -- i.e., the mechanism itself. And you also mingle the straight observations with subtle assertions about invisible causal links.

> To say that histories "change organisms organically" is to say precisely nothing, because nothing is said about WHAT organic changes, with what significance, occur. It's like explaining the action of a drug on the body by saying that it's "chemical."

As I said, the statements and restatements, and other sayings don't mean anything to anyone other than those committed to statements and sayings of that particular sort.

>> To Skinner, natural selection is a basic law (causal mode, even) of biological evolution, the ontogeny of individual behavior, and cultural practices.

> Natural selection is not a basic law, it's a way of characterizing in a couple of words a whole interaction between a species and its environment. The use of the word "selection" is completely metaphorical; taken literally, it's an effect without a cause (that is, selection occurs, but there is no selector).

That's partly why I said 'natural selection' doesn't explain anything. [Sure, natural selection is a step up from supernatural selection, and it does appear that we have folks today who haven't yet taken even this step.]

>> Bill, you didn't comment on my suggestion that PCT modelers base models on PCT units not on units from non-PCT areas such as operant psychology.

> We would be happy to do so, if operant psychologists didn't keep insisting that they have already explained the same phenomena we explain, and if their explanations of phenomena we haven't yet explained were credible.

I'm not sure how this follows from my suggestion for a direction for PCT modeling.

> Among proponents of Skinner's interpretations, I see nothing in the way of seeking independent confirmation. Instead what I see is a continual effort to show that the proffered explanations would indeed predict what we observe. If some difficulty arises, like data that don't seem to jibe with the basic picture, the reaction is simply to start manipulating the basic variables of reinforcement theory until a combination is found that will lead to at least a qualitative approximation of what is observed. Since in most situations many variables can be identified, showing many different relationships to behavior, one can always find a variable to play the role of the reinforcer or the discriminative stimulus, and that is usually sufficient to construct a plausible explanation -- one that fits what was observed.

> What I don't see is any attempt to challenge the theory -- to ask how we could verify by some means other than pointing to the phenomenon to be explained that a reinforcing effect occurred, or that a discriminative stimulus was in fact detected by the nervous system of the organism. The general attitude I find is that no other interpretation of the observations is even conceivable within a scientific framework.

Wrong--you'd be surprised at the different views that fly under the banner of operant psychology, behavior analysis, EAB, or whatever label--they can't agree on even this. Apparently, some of those who do not follow this literature closely might find an overview of all this illuminating. I know of no such overview in print. But oh there have been challenges. Read the Skinner-Herrnstein papers in the _American Psychologist_ (1979?, a guess). It appears that Skinner saw Herrnstein as his successor at Harvard. Well, Herrnstein stayed at Harvard, but strayed so far from orthodoxy that he became a 'bad psychologist.' I am referring to this vast amount of EAB work in molar behaviorism--matching law, correlation-based law of effect, cognitive constructs--all so much un-Skinnerian yet firmly in "his" branch of behavioral science. A recent example of this is the last book Howard Rachlin published--Rachlin is considered a leader in EAB, right? But the book (I have a review of it in a recent issue of the _Psychological Record_) absolutely infuriated three of my colleagues who are rather devoutly Skinnerian (even though they may not admit to this directly). They are basically out of the Kansas sub-cult of behavior analysis-- dust-bowl behaviorists (Don't me hear any of that stuff about under the skin unless you're talking motoric movements in space). Rachlin, by the way seeks to reconcile behavior analysis and cognitive/physiology, a sure way to talk to raise the hackles of a good Skinnerian. One more item on challenges to Skinner. Another hot bed of Skinnerian psychology has been Western Michigan University. For years, all doctoral students from there were well-trained Skinnerian psychologists--there was no other way they could get out. Linda Parrott (now Linda Hayes) was one of these Ph.D's. I know of no one who has challenged Skinner's theory with more authority than has Linda. Throughout all of her attacks, Linda has been very active in the Association for Behavior Analysis, an organization Skinner seemed to consider "his." So, there have been many challenges to operant theory from within. The much acclaimed 'attack' of Chomsky's is pathetic against some of the challenges from Skinner's students themselves and from students of Skinner's students (especially students of Herrnstein). I seems that someone ought to put together all the challenges to Skinner's ideas that arose from his own camp. There is reason to believe that Skinner was quite aware of all this, so much so that he published an infamous letter that embarrassed many of his troops in his last year or so.

Dennis Delprato

Date: Wed, 21 Jun 1995 10:24:19 -0600
Subject: Operant psychology and PCT

[From Bill Powers (950621.0730 MDT)]

Dennis Delprato (950620) --

I do appreciate your thoughtful post on Skinnerian psychology and its successors.

> It seems that someone ought to put together all the challenges to Skinner's ideas that arose from his own camp.

This, as well as a sympathetic but firm discussion of the whole field of behavior analysis and Skinnerian theory, would make a book of outstanding importance. There are few people I know about -- you and Bruce Abbott are the only two who come to mind -- who have had a deep involvement in the Skinnerian movement and also have the advantage of an outside viewpoint other than that of conventional psychology from which to speak. You could correct many misconceptions about radical behaviorism, and at the same time work toward leading that field into a more advanced conception of modeling behavior. Isn't it about time to do this?

> They use 'history' in a temporal sense--what happened previously in time. If you are saying that what happens to an organism historically is a function of the organism's behavior, operant theorists would agree.

This is partially what I mean. The rest of what I mean has to do with the _interpretation_ of the role of historical events. There is far more involved in the Skinnerian interpretation than a mere listing of spatiotemporal processes over a period of time. There is a deep-seated _causal interpretation_ that has the same roots as do other ideas in psychology.

The whole thrust of psychology as a new science was to interpret the behavior of organisms as the behavior of a mechanism like any other physical mechanism. I'm sure this orientation was consciously adopted very early in the history of psychology, as a way of catching up with physics, the Big Brother of science. It was adopted also as a counterthrust against what scientists have always regarded as the enemy of progress: imprecision, superstition, mysticism, and magic. During all the formative years of scientific psychology, there was no known alternative to the physical-science approach that did not in fact entail some form of mysticism and magic.

All this began to change in the middle of the 20th Century when the automation revolution and the computer revolution began. Old concepts of mechanism began to crumble as amazing discoveries were made about the capacities of artificial computing devices and control mechanisms. The firm boundary between "mental" and "physical" phenomena began to waver, until now nobody who is knowledgeable about the new machines takes the distinction seriously.

This is one reason, Dennis, that I see your firm insistence on "spatiotemporal" explanations as somewhat old-fashioned. It has the flavor of the old conception of mechanism, in which mental phenomena are considered ghosts in the machine rather than realities of the physical universe. What we have discovered about the capabilities of artificial computers and control systems has taught us that organization and function have a reality just as certain as the reality of matter and energy. Not only that, we have found that organization and function, while residing in physical matter, can transcend physical matter in the sense that the same organizations and functions can be realized in an endless variety of physical instantiations. The particular physical particles and physical laws that are used to implement computing or control functions are unimportant; every year more and different physical means are found for storing bits and performing logical functions. What matters is how the various physical elements are _related_ to each other, not what the physical elements are. And these relationships are not spatiotemporal in nature. They are functional and organizational -- what used to be called "mental."

Arthur C. Clark was the one who said "The products of any sufficiently advanced civilization would seem to us to work by magic." This is how the models of any sufficiently advanced theory of human behavior would appear to a conventional scientific psychologist, even a "radical" one. A great part of Skinner's emphasis was on showing the superficiality of mental explanations of behavior, explanations that invoked unobservable entities and intervening variables. For every mental explanation, he tried to supply an equivalent explanation that involved only externally observable relationships among physical variables. Among the explanations he sought to displace were those that relied on attributing purpose and intention to the organism. He saw these as simply more intervening variables in the same class as traits, dispositions, expectancies, aspirations, fears, and thoughts. As far as he could possibly have known, all these concepts were just fictions, unprovable and untestable by their very nature. Skinner could not distinguish between a bad model of brain function and a good model, because he had never seen a good model.

Skinner's great mistake, which he was far from the only one or the first one to commit, was to assume that everything necessary to give a full and accurate account of behavior could be found in the external circumstances surrounding an organism. He assumed that if he just recorded very carefully what happened, he would be on firm ground because anyone, at any time in the future, could record the same events and verify his analysis. This is the faith of the empiricist; that a careful, honest, and well-trained observer cannot make any substantial mistakes. It is also the faith of the naive realist, who would be extremely surprised and skeptical about any claim that others, equally well-trained and equally honest, would not necessarily see exactly what he sees.

Remember my example of the lawn ornament on which a little wooden man busily turns a crank that turns the vanes of a windmill to cause the wind to blow. In the absence of any reason to doubt the interpretation, one could trace cause and effect point by point through the linkages, and show that this interpretation makes perfect sense. One could even show that when circumstances changed, the same interpretation still worked: when the little man stops cranking, the wind stops, too.

We don't make this mistake with the lawn ornament because of a vast network of understanding and observational detail that surrounds our observations of it. But when that vast network is missing, as it has been for the behavior of organisms, we can easily think we are simply reporting a causal relationship when we have it exactly backward. That is what happened to Skinner with the concept of reinforcement.

He was almost doomed to make this mistake because of a strong prior assumption: that all behavior was caused, ultimately, by the environment. A reinforcer and a change in behavior are two observable events. If the environment causes behavior, then clearly the reinforcer must cause the change in the behavior. The basic assumption permits no other interpretation even if one is inclined to make it. So what Skinner reported as a trained observer was exactly what he was trained to believe.

The most frustrating (to me) aspect of Skinner's early work was not this interpretation, which was really standard among scientific psychologists, but the fact that he then proceeded to set up an apparatus in which a reinforcer would never appear unless the behavior occurred first. The sequence of events always began with a behavior, followed by a reinforcement. This causal relationship was right there in the environment, visible even to an untrained observer.

Once an experiment got under way, however, reinforcements and behaviors occurred in a continuous stream, so as with the lawn ornament one could choose either way of perceiving causes and effects. As the animal worked toward mastery of the task, one could see the gradual increase in reinforcement rate as causing a gradual increase in behavior rate, or the gradual increase in behavior rate as causing a gradual increase in reinforcement rate. Once the rates had leveled out, one could see either the reinforcements maintaining the behavior, or the behavior maintaining the reinforcements. In the absence of any larger network of understanding of behavior, either interpretation would fit the observations.

Skinner, unfortunately, understood that the causes of behavior always lie in the environment. He therefore had no option but to see the reinforcement as the cause and the behavior as the effect, both during acquisition of the behavior and during maintenance of the behavior.

Along with the basic method came the so-called schedule of reinforcements, implemented by an automatic apparatus or a human being working strictly to a rule. The role of this schedule could also be interpreted in two ways. It could be said that the behavior, acting as input to the schedule, caused an output in the form of periodic reinforcements. Or it could be said that the contingency embodied in the schedule, in combination with the effects of the reinforcements, caused the pattern of behavior. Once again, Skinner's basic assumption left him no choice as to how to report his observations. Clearly the contingency and the reinforcement combined to control the behavior. The trained observer once again observed what he had trained himself to see.

All subsequent interpretations of experimental results were required to remain consistent with the initial Gestalt resulting from the choice between the two possible directions of causation. If the initial choice was wrong, then necessarily all subsidiary and subsequent interpretations had to be wrong, too. The numerical records of reinforcement and behavior rates, of contingency ratios and combinations of ratios, would remain quite accurate, but in every instance their interpretation would depend critically on the initial interpretation of the relationships among behaviors, contingencies, and reinforcements.

The integrity of radical behaviorism, therefore, rests on the correctness of Skinner's initial interpretation. Every conclusion about behavior drawn by every subsequent radical behaviorist as a result of every subsequent experiment depends for its validity on whether Skinner saw the lawn ornament in the right way. If Skinner got it backward, then so has everyone who followed him, from the very beginning to this very moment.

I think that in his or her heart of hearts, every radical behaviorist who has come up against PCT knows that this is the basic threat of PCT. For PCT does not portray reinforcers as causes and behaviors as effects. Quite the contrary: PCT shows behavior as the means by which the organism acts on its environment to produce consequences that are to its liking. If the reinforcer is not there, the organism finds a way to bring it into being. If there is not enough reinforcer, the organism changes its behavior to make more of it. If there is too much reinforcer, the organism acts to reduce it. It is the organism that determines how much reinforcer there is to be, and that learns to create whatever action is required to produce that much reinforcer. In fact, the name "reinforcer" totally loses its meaning as something that can make the organism do something. The reinforcer is an effect, not a cause.

So if Perceptual Control Theory is right, Radical Behaviorism is wrong. It is as simple as that. The fundamental assumptions about the direction of causality are diametrically opposite; there is no reconciliation possible, no meeting in the middle, no compromise that will satisfy both theoretical structures. As with the lawn ornament, one must choose an initial interpretation; from that initial interpretation _all else follows_.

So, Dennis, the historical record of which you speak does not speak for itself. It must be interpreted, and how it is interpreted makes all the difference. I think I have strong evidence, incontrovertible evidence, that Skinner chose the wrong interpretation and sent thousands of his followers down the wrong path. To correct this error it is necessary to go back to the initial bifurcation and see what happens if we start with a different interpretation of the same events.

Best to all, Bill P.

Date: Wed, 21 Jun 1995 14:16:28 -0700
Subject: A Problem for Reinforcement Theory?

[From Rick Marken (950621.1410)]

The following is data from the R. coli experiment. It shows a running average of cursor position over time.

```
                                     x = random
        y                            o = reinforcement
                                     z = subject
        |
  start |xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
        | o
        |  ozzzzzzzzzzzzzzzzzzzzzz
  target|   o ooooooooooooooooooooooooooooooooooooooooooooo
        |                          z
        |                           z
        |                            zzzzzzzzzzzzzzzzzzzzzz
        |
        |
        ----------------------------------------------------
                                  Time
```

The x's are the average cursor positions that result from random responding (probability of response is .5 at each time instant); the cursor stays near the starting position (start), which is about 75 pixels above the target position (target).

The o's are the average cursor positions that result from running a reinforcement model; this model is a pure learning model (no extinction) so P(R|S+) is 0 and P(R|S-) quickly goes to 1.0. The model responds when and only when the cursor moves away from the target. The model quickly moves the cursor to the target position and keeps it there.

The z's are the average cursor positions that result from running a human subject. Like the reinforcement model, the subject quickly moves the cursor to the target (the z and o lines actually overlap though the first half of the graph). But during the middle of the run the cursor quickly moves to a new position and stays there. This behavior of the subject is in marked contrast to that of the reinforcement model.

Here's a question for EAB experts: How do I change the reinforcement model to make it behave like the subject?

The reinforcement model was very simple. I just incremented P(R|S-) according to:

$$P(R|S-)_{n+1} = P(R|S-)_{n} + .8\ (1-P(R|S-))$$

when R occurred in the presence of S- and the result was a reinforcement (movement toward the target). P(R| S-) was NOT decremented when a press in the presence of S- was not reinforced. P(R|S+) remained at 0; the model never pressed when the cursor was moving toward the target (S+).

What seems to be happening with the subject is that, in the middle of the run, P(R|S-) suddenly changes; the probability of a press when the cursor is moving away from the target, P(R|S-), becomes much less than 1.0. And P(R|S+), the probability of pressing when the cursor is moving toward the target, suddenly becomes much greater than 0.

How does reinforcement theory explain this change in P(R|S-) and P(R|S+)?

Best Rick

Date: Thu, 22 Jun 1995 17:42:29 -0500
Subject: Back in the Saddle Again

[From Bruce Abbott (950622.1740 EST)]

Well I'm back. Actually I've been back since yesterday but something fried in our network and I've had no ability to send anything until just now, although I have been able to catch up on my CSG-L reading to some extent. I'll try to organize a reply to some of the reinforcement stuff when I can; right now I'm facing a deadline on some textbook work.

Meanwhile:

Rick, I don't understand your new R. Coli demo, or rather, I don't understand why your "subject" should show the behavior you diagrammed (950621.1410).

> The z's are the average cursor positions that result from running a human subject. Like the reinforcement model, the subject quickly moves the cursor to the target (the z and o lines actually overlap though the first half of the graph).

O.K. so far...

> But during the middle of the run the cursor quickly moves to a new position and stays there.

Why? Shouldn't the subject be trying to keep the cursor on the target?

> Bill Powers (950619.1455 MDT)

> Some thoughts on reinforcement theory that must have passed through the minds of those who developed it, if not exactly in the order presented below. And following, a possible resolution of the conflict between reinforcement theory and control theory.

An interesting post that deserves extended comment--I'll try to get that done fairly soon...

My "vacation" trip took me to Fredericktown in Knox Co., Ohio to do a little genealogical research. Looking up land sales at the courthouse in Mt. Vernon, I found the following record regarding a great-great-grandfather of mine, Thomas Barrington:

Thomas Barrington
DEED to
John Powers
Recd Decr 6, 1851

Bill, could John be related to you?

Regards, Bruce

Date: Thu, 22 Jun 1995 19:25:00 -0500
Subject: Re: Reinforcement theory: As you like it

[From Bruce Abbott (950622.1920 EST)]

>Rick Marken (950620.1030) --

> I think it is significant that reinforcement theorists must develop a new version of their theory each time a new behavioral situation is described. For example, in a simple operant conditioning experiment, the occurrence of food following a bar press is considered a reinforcement. I have never heard reinforcement described as the change from "no food" to "food". So, when I developed the E. coli demo it seemed reasonable to assume that the direction of movement following a press was the reinforcement; like the food in the basic operant conditioning experiment, direction of movement is a direct result of a response. But, in order to make the reinforcement

model of E. coli behavior work, it was necessary to redefine reinforcement as _change_ in the direction of movement after a press; change toward the target was the reinforcement.

> It is also sometimes necessary to change what was considered a non-reinforcer or punisher in certain experiments. This is what happened in the first version of my R. coli demo. In that demo, a press leading to a change from non-target to target is considered a reinforcer but a change from one non- target to another could not be considered a punishment because, if it were, the reinforcement model wouldn't work.

What nonsense!

It's like saying PCT can't be right because you have to write a different model to account for behavior in the e. coli task and standard pursuit tracking. To paraphrase, "I think it is significant that PCT theorists must develop a new version of their theory each time a new behavioral situation is described." The definition of the controlled variable keeps changing--here it was nutrient level and there it was difference between target and cursor position. What changes is not the theory but the model which applies the theory to the situation in question.

> I have never heard reinforcement described as the change from "no food" to "food".

Well, your understanding of reinforcement theory is about 25 years out of date. For example, see Baum, W. (1973). The correlation-based law of effect. _JEAB_, _20_, 137-153.

> But, in order to make the reinforcement model of E. coli behavior work, it was necessary to redefine reinforcement as _change_ in the direction of movement after a press; change toward the target was the reinforcement.

This sounds like a recitation of history but actually it is simply a product of your imagination.

How can you have an "event" without change? Reinforcement _always_ involves a change of condition. Reinforcement was not "redefined," it's the same as it's always been.

> In that demo, a press leading to a change from non-target to target is considered a reinforcer but a change from one non- target to another could not be considered a punishment because, if it were, the reinforcement model wouldn't work.

No, not only must there be change, but to be reinforcing the change must lead to improvement with respect to the desired condition. Responses that lead to irrelevant beeps and clicks aren't punished, they simply go unreinforced. In contrast, having the rate of increase in nutrient deteriorate following a tumble definitely represents a worsening of conditions and thus would be expected to punish the tumble.

> The events that are considered reinforcement, punishments, etc must be changed to account for behaviors in different situations because the behavior of the reinforcement model is determined by the situation. The model will only work (account for the actual behavior observed) if the modeler can pick out aspects of the environment that will produce the right behavior. This, of course, means that the reinforcement model cannot fail as long as one can find appropriate environmental inputs (appropriate consequences of responding) for the model.

This just restates your erroneous thesis. But you do bring up a point. ANY model based on theory involves speculation about what the relevant variables are in a given situation. In developing a control model, for example, you must first decide what the controlled variable is. Your guess might be wrong, and you will then have to try another. Does this mean that control theory is wrong? Of course not. Same here: In applying reinforcement theory to a specific situation, I am speculating at to what variables may involved, and naturally I'm going to pick variables that allow the model to perform correctly (don't you?). Whether we're talking reinforcement theory or PCT, the crucial step is to TEST the assumptions of the model, i.e., perform tests to determine what the actual reinforcer is, or the actual controlled variable. Our little modeling efforts have not included that critical step, so the models remain speculative. The fact that we have not gone beyond speculation is not a criticism of the theory on which a model is based, reinforcement OR PCT.

> Of course, the power of reinforcement theory exposes its basic flaw. The power of reinforcement theory is achieved by looking for aspects of the environment that can account for the behavior we see. But this means that the same aspect of the environment must be treated as a reinforcer in some circumstances and a non-reinforcer or punisher in others. Since the theory provide no independent means of measuring the reinforcingness or punishingness of the environment, we are free to ascribe these changing characteristics to the environment as necessary to explain behavior. This is not science; it is tautology.

Based as it is on a misconception of how reinforcers and punishers are defined, and of what we are doing with these models, this conclusion is moot. How about "since the theory (PCT) provides no independent means of measuring the controlled variable, we are free to ascribe these changing characteristics to the environment as necessary to explain behavior." Your statement regarding reinforcement and punishment is as true as my paraphrase, and for the same reason. It sure sounds good.

Regards, Bruce

Date: Thu, 22 Jun 1995 20:20:04 -0700
Subject: Re: Reinforcement Theory

[From Rick Marken (950623.2020)]

Bruce Abbott (950622.1740 EST) --

> Rick, I don't understand your new R. Coli demo, or rather, I don't understand why your "subject" should show the behavior you diagrammed (950621.1410).

I don't either. I presume it is because this behavior was selected by its consequences. But I'm having trouble seeing how this might have happened. What do you make of it?

Me:

> I think it is significant that reinforcement theorists must develop a new version of their theory each time a new behavioral situation is described...I have never heard reinforcement described as the change from "no food" to"food".

Bruce Abbott (950622.1920 EST) --

> What nonsense! ... What changes is not the theory but the model which applies the theory to the situation in question.

I seem to recall that a reinforcer was defined as a consequence of responding that increases the probability of the response that produces it. There is nothing in that definition about the consequence being a _change_ from what it was before the response to what it is after.

In E. coli I empirically determined the reinforcing value of each consequence of a response based on the notion that the reinforcingness of a consequence could be measured by the probability, after the consequence, of the response that produced it; if the consequence of a response was movement away from the target (regardless of what the movement relative to the target had been before the response) the probability of a response after this consequence was high (the interval between responses was brief); if the consequence of a response was movement toward the target (again, regardless of what the movement relative to the target had been before the press) the probability of a response after this consequence was low (the interval between responses was long).

> your understanding of reinforcement theory is about 25 years out of date.

I'm willing to believe that and that's why I accepted your explanation.

> In developing a control model, for example, you must first decide what the controlled variable is.

This is not true. When we develop a control model we know what the controlled variable is; otherwise there is no reason to develop a control model in the first place.

> Your guess might be wrong, and you will then have to try another. Does this mean that control theory is wrong?

No. It means that the person trying to apply control theory doesn't understand what the theory is about. Control theory is about control; you can't even sensibly talk about control until you know what variable(s) is (are) being controlled.

Me:

> Since [reinforcement] theory provides no independent means of measuring the reinforcingness or punishingness of the environment, we are free to ascribe these changing characteristics to the environment as necessary to explain behavior.

Bruce:

> Based as it is on a misconception of how reinforcers and punishers are defined, and of what we are doing with these models, this conclusion is moot. How about "since the theory (PCT) provides no independent means of measuring the controlled variable, we are free to ascribe these changing characteristics to the environment as necessary to explain behavior." Your statement regarding reinforcement and punishment is as true as my paraphrase, and for the same reason. It sure sounds good.

Actually, as you well know, there IS an independent means of measuring the controlled variable; it is The Test for the controlled variable. The Test is of critical importance to PCT. If you don't know that a variable is under control or what that variable is, then you have no reason to try to model behavior using PCT.

Actually, I think that the "old fashioned" definition of a reinforcement (as a consequence that increases the probability of the response that produced it) does provide an independent means of measuring the "reinforcingness" of environmental events. But, as you say, my understanding of reinforcement theory is about 25 years out of date (well, 15 years; I did read Skinner's 1981 Science article) so maybe now there is no independent means of measuring reinforcingness at all.

Best Rick

Date: Fri, 23 Jun 1995 07:16:26 -0500
Subject: Re: Reinforcement Theory

[From Bruce Abbott (950623.0715 EST)]

>Rick Marken (950623.2020)

>>Bruce Abbott (950622.1740 EST)

>> Rick, I don't understand your new R. Coli demo, or rather, I don't understand why your "subject" should show the behavior you diagrammed (950621.1410).

> I don't either. I presume it is because this behavior was selected by its consequences. But I'm having trouble seeing how this might have happened. What do you make of it?

It's a control system that doesn't control? Obviously I'm missing something here. Help me out.

> I seem to recall that a reinforcer was defined as a consequence of responding that increases the probability of the response that produces it. There is nothing in that definition about the consequence being a _change_ from what it was before the response to what it is after.

I can't imagine a "consequence" without a change, can you? If the "consequence" of a response is that nothing changes, then the response has no consequence. Thus if your cursor is drifting to left, away from target, you press the button, and nothing changes, the response has had no consequence and has been neither punished nor reinforced. (Note: I am ignoring possible "response cost" here, which I assume is slight.)

>> In developing a control model, for example, you must first decide what the controlled variable is.

> This is not true. When we develop a control model we know what the controlled variable is; otherwise there is no reason to develop a control model in the first place.

That would depend on the available information. For example, I seem to recall your assisting me in developing TWO models of control in the inverted-t illusion situation, one based on a difference between two perceived lengths and the other on their ratio. Seems to me we didn't know what the controlled variable was until we developed the models and then determined empirically which one generated a better fit to the data. Furthermore, even those two possibilities were only (informed) guesses. The actual variable being controlled might be something different yet, like a weighted combination of difference and length. So don't tell me you always "know" what the controlled variable is when you develop a control model. It isn't true.

>> Your guess might be wrong, and you will then have to try another. Does this mean that control theory is wrong?

> No. It means that the person trying to apply control theory doesn't understand what the theory is about. Control theory is about control; you can't even sensibly talk about control until you know what variable(s) is (are) being controlled.

I see. So now you're telling me that you developed two models of control in the inverted-t task because you didn't understand what control theory is about. This is getting interesting...

> Actually, as you well know, there IS an independent means of measuring the controlled variable; it is The Test for the controlled variable.

Yes, and there is a similar independent means of determining whether a consequence is a reinforcer, punisher, or neutral event under given conditions. Thorndike described it in 1898.

Regards, Bruce

Date: Fri, 23 Jun 1995 08:58:08 -0700
Subject: Re: Reinforcement Theory

[From Peter Burke (950623 9:30 PDT)]
>[Rick Marken (950623.2020)]

> I seem to recall that a reinforcer was defined as a consequence of responding that increases the probability of the response that produces it.

If a reinforcement is defined as a consequence of responding that decreases the error between the perception and the reference signal then there would seem to be little difference between PCT and reinforcement theories!! ;-)

Peter J. Burke

Date: Fri, 23 Jun 1995 10:19:10 -0600
Subject: Re: reinf theory

[From Bill Powers (950623.0730 MDT)]

Bruce Abbott (950622.1740 EST) --

Welcome back.

What Rick is trying to do (beneath all the din) is find a situation that will distinguish between the reinforcement model and the control model. I don't think he's done it yet -- it's necessary to equalize the probability that a behavior will be rewarding or punishing, without making control impossible, too.

I was able to modify the properties of the tumbling in the E. coli model so that no matter what direction E. coli is traveling, the probability of a tumble improving the rate of change of nutrient is equal to the probability of making it less favorable. This drastically reduces the success rate of your basic Ecoli4 model. I say drastically reduces rather than eliminates because by chance it is possible that the probabilities will depart from equality in the direction that produces some control, so the goal may be reached eventually. In such cases we find that the four frequencies of outcomes were in fact not equal. On many trials, of course, the spot disappears from the screen and never (i.e., within five or ten minutes) comes back. With this setup, a human operator can still get to the goal every time, but the model can't.

All this shows is that in this special environment, the control model works where the reinforcement model, as given, doesn't. But I can't show that NO definition of a reinforcer and discriminative stimulus would work.

The only final answer will come from comparing the details of real behavior with the details of what the model predicts. For example, when you vary the delay between tumbles by using a probability calculation, the relationship between delay and angle of travel will have a rather large random component; there should be a calculable proportion of long delays, for example, while traveling away from the goal. In a model that uses a systematic output function, this would not happen. When the probabilities reach 0 and 1, however, this difference would disappear. Other possibly observable differences would be in the apparent reference setting, and in the slope connecting error to delay.

I don't think, however, that we really want to spend a lot more time on E. coli. If it takes this much work to distinguish between two theories, the results will never be clear-cut. And we're not really comparing the two models correctly anyway, because the control model is a performance model, not a learning model. To test the reinforcement theory of learning, we would have to propose _other learning models_ and test them against the reinforcement model. We should be comparing the reinforcement model with a Kalman Filter model, or a random reorganization model, or some such model. With E. coli that would get very confusing, because the behavior that is to be learned also uses the E. coli method.

So far, we've been playing the game using the conventional rules: constant reference signal, no disturbance. I think this is really the reason for the difficulty in distinguishing the two models. Even just introducing disturbances of the controlled variable should show up some major differences. A change of reference signal would also cause problems for reinforcement theory: a sudden change in the definition of the reinforcer, for no externally-explainable reason (is this a hint about the data Rick presented?).

We could make such a comparison in the rat experiments that will be coming up pretty soon now, I hope. If we start with a simple fixed ratio schedule, we could develop a control model and a reinforcement model that would fit the behavior over a range of schedules.

Once the behavior is well-learned, we would be talking about a performance model (in the PCT case at least). According to the PCT model, slowly but unpredictably changing the ratio should lead to a change in behavior rate in a specific relationship to the varying ratio, with the obtained reinforcement rate varying considerably less than the behavior rate does (and in the _opposite_ direction). In fact, from the model fitted to the data over the range of schedules, we would be able to predict how much the reinforcement rate and behavior rate would change as the ratio is varied. For example, we could start with a baseline ratio of 6, and vary it slowly and randomly in the range from 1 to 11 (five ratios higher than 6, and five lower). We would then compare the behavior of a simulated control model with fixed parameters with the behavior of the real rat (we would fit the control model to each rat, not to the average over rats). Of course we would do the same using the reinforcement model you develop.

Another way of introducing disturbances would be to add and subtract reinforcements arbitrarily. This, however, would be difficult to distinguish from varying the ratio.

With respect to my proposal of yesterday, by keeping track of when the rat is actually at the bar and when it is elsewhere, we might be able to test the idea that the reinforcement model applies most naturally to the process of acquiring the right kind of behavior rather than the right amount of a single kind of behavior.

As to John Powers, I don't know much about my own family's history. As I recall, my Powerses passed through Ohio and Illinois before striking out to homestead in the Far West (my father spent his youngest years on a homestead in Christmas Lake Valley, Eastern Oregon). It's possible that there's a connection. If my ancestor didn't finish paying for that deed from your ancestor, I may owe you some money. Sounds like your ancestor settled down while mine moved on: typical Powers restlessness.

Best, Bill P.

Date: Fri, 23 Jun 1995 11:39:05 -0500
Subject: Re: The Domains of Reinforcement Theory and PCT

[From Bruce Abbott (950623.1135 EST)]

>Bill Powers (950619.1455 MDT) --

> Suppose we thought of B and R as having variable magnitudes. Now we can have more or less of B, and more or less of R. This allows R to have an effect on B in terms of magnitude. If we assumed that an increase in R causes an increase in B and that R is proportional to B, then we would have

> R1 = k1*B1
> B2 = B1 + k2*R1.
> R2 = k1*B2
> B3 = B2 + k2*R2,

> and so on. From this we would quickly conclude that the behavior B will either get larger and larger without bounds (k1*k2 1) or smaller and smaller until it is zero (k1*k2 < 1).

> This model has no stable states other than infinity or zero. This is not what we want as a model of learning, either.

It is easy to modify the model, however, so that it no longer tends to infinity. For example, think of reinforcement as being equivalent to charging a capacitor.

In this case the increment in response strength is proportional to the difference between its present strength and some maximum possible strength.

> What we can do is go back to the idea that B and R are events that occur or don't occur, but consider the _probability_ that B will occur. Now we can say

> B --> R (the behavior, if it occurs, leads to R)

> pr{B} = pr{B} + k*R {where R is either 1 or 0)

> This arrangement does not cause behavior to run away to infinity, because now the greatest possible probability is 1. If the behavior starts with a zero probability of occurring and the probability is increased every time R occurs, eventually the probability of B occurring will become 1, and B will occur every time R occurs.

> This, then, gives us a basic model of learning with the right properties. At first there is no behavior or very little. When the behavior does appear, it produces a reinforcement, which in turn increases the probability that the same behavior will occur again. With enough repetition, the behavior will eventually occur every time the reinforcer occurs. Since the behavior always results in a reinforcement, the behavior will persist.

Yes, if we wish to view behavior as a series of well-defined "responses" such as pressing a lever, we can then talk about the rate over time at which these responses occur OR in terms of the probability of the response. The probability analysis provides natural limits of 0 and 1; the rate measure provides a natural lower limit of 0 but the upper limit depends on further considerations such as minimum interresponse time and level of motivation. I point this out merely to indicate that one is not necessarily driven to a probability analysis as a means of avoiding a runaway to infinity.

> This model can now be expanded to include a discriminative stimulus S, which signals the conditions under which B will produce R. Now the probability that is altered by R is Pr{B|S}: the probability that occurrence of the discriminative stimulus S will lead to behavior B. Again there is no runaway condition, because the maximum possible probability of any event is 1. The most that can happen even with continued reinforcement is that every time S occurs, B will occur.

The discriminative stimulus should not be viewed as a stimulus that elicits a response. Pr(B|S) is not the probability that S will lead to B but the probability density over time of B, given that S is present. If this probability is high and S is short, the result may look like S-R: S occurs and R follows almost immediately. But this is only an extreme case.

> If the probability in question is considered to be a probability _density_ (the probability of occurrence per unit time), the probability of B can be interpreted as the average interval between occurrences of B. If the probability density is high, B will occur after only a few time-increments; if low, B will occur only after many time increments. This leads to seeing the natural measure of behavior as _rate of repetition_, where the average rate of repetition is simply the reciprocal of the probability density.

Yes, and Pr(B|S) can be interpreted as the rate of repetition of B given S.

> Thus even though behaviors and reinforcements themselves are considered as unitary events which either occur or don't occur, we can find a continuous variable in terms of which to measure behavior: its rate of occurrence, which is closely related to the basic measure of probability of occurrence per unit time. The rate measure of behavior and reinforcement is thus the natural result of considering behaviors and reinforcements to be events, and of proposing that the effect of a reinforcement on the production of a behavior (or on the response to a discriminative stimulus) occurs via an effect on the probability-density of the responses. The use of probabilities is dictated by the fact that if the magnitude of

reinforcement had an effect on the magnitude of behavior, the resulting model would have stable states only at zero and infinity.

Again, the use of probabilities is not "dictated" at all, as one may work directly with the rate measure. In cases where what varies is the magnitude of the response (i.e., response force), the reinforcement principle has been applied successfully without recourse to a probability interpretation. This is why reinforcement theorists talk about reinforcement affecting the "strength" of behavior rather than its probability. Rate or probability are considered to be measures of response strength, but so would the force exerted on the lever if the reinforcement contingency required a specific force before delivery of the reinforcer.

> We can see that as originally conceived, the model is a positive feedback model. The more reinforcement there is the more behavior there is, the more behavior there is the more reinforcement there is. The converse also holds true; if there is any lessening of either behavior or reinforcement, both variables must decrease with the only limit being zero.

As each response continues to produce the "same" reinforcement, the rate of behavior will increase to some value dependent on the value of the reinforcer to the subject and the cost of responding (e.g., effort). The "amount" of reinforcement conceived as reinforcement per response is constant in this scenario and this relationship determines how much response "strength" is incremented with each successive reinforcement, up to its limit.

This "molecular" analysis based on the immediate consequence of each response can be supplemented with a "molar" approach based on rates if one takes the view that the subjects in these experiments can integrate events over time. In this view, receiving individual reinforcements at a higher rate is more reinforcing than receiving them at a lower rate. A brief increase in response rate (e.g., accidental lever presses as byproducts of general activity) result in an increase in reinforcement rate (increased rates of lever-pressing are reinforced by increased rates of food pellet delivery). Decreased rates of lever-pressing lead to decreased rates of food pellet delivery (punishment). The molar approach thus predicts that response rates will tend to increase toward the maximum sustainable rate as determined by reinforcer value and response cost.

This is what is observed, and not the bi-stable behavior you predict on the basis of the positive feedback relationship as you conceive it. The analysis you propose would lead to instability of behavior on a ratio schedule: any decrease in response rate (perhaps resulting from the rat's momentary need to scratch an itch) would be expected to lead to total extinction of the response. It doesn't happen, and it isn't predicted to happen by reinforcement theory.

> To avoid this obviously inappropriate result, the meaning of "more" had to be modified so that the outcome was not a runaway condition. This is the function of the concept of increasing _probabilities_ rather than _magnitudes_. The model remains a positive feedback model, but the insertion of a probability puts a limit on the runaway condition where the probability becomes 1.

But this doesn't solve the problem, does it? Your analysis still has the instability problem I pointed out whether "more" is an increase in rate or an increase in probability.

> Let's go back to the original model, but this time add a reference signal and change the system to a negative feedback model:

```
>                          R'*
>                           |+
>                       --------
>              ------>|  ORG   |---->--
>             |        --------        |
>             |                        |
>             |                        |
>             |         --------       |
>             R<----- |  ENV   |<------B
>                       --------
```

[This was followed by a description of the control model.]

Let's apply this model to the steady-state situation involved in performance on a simple CRF (1 response per reinforcement) schedule as ordinarily studied in the operant chamber. A hungry rat is placed in the chamber. We might conceive of the set point for rate of food pellet consumption as, say, 30 pellets per minute (30 ppm) under these conditions (essentially continuous eating). However, the apparatus limits the maximum rate to 10 ppm because of the delays involved in depressing the lever, moving to the food cup, picking up the food, devouring it, and returning to the lever. Thus, there is NO WAY that the rat can reach its reference level for this quantity, although it can reduce the error by a significant amount through lever-pressing (and by no other means). I believe that control theory predicts that the rat (once it has learned what to do) will develop a rate of responding that minimizes the error, up to the point where the error is reduced enough to bring the output below maximum. Depending on the gain, there will be a region within which the rate of responding will be a direct function of the magnitude of the error.

If we now increase the ratio requirement, the maximum rate of reinforcement available on the schedule is less (say, 5 ppm). This means that a larger error remains after responding has reached its maximum rate than was the case with the CRF schedule. If response rate on the CRF schedule was already at maximum, it will still be at maximum on the new schedule; if not, then it will increase.

But this analysis ignores response cost. Assume there is a second control system for "response effort" and that the reference for this perception is set to a low value. We now have two control systems in conflict when the rate-of-eating system brings the rate of lever-pressing up and thereby brings response effort above its set point. The result is that neither system can reach set point and rate of responding stabilizes at a kind of equilibrium value between the two set points. As the ratio requirement increases, the rate of responding increases, but not enough to keep the rate of food intake constant; instead the latter declines somewhat with increasing ratio requirements. You get the right-hand portion of the inverted U function shown by Motheral's data.

A similar analysis in terms of reinforcement (response-contingent food delivery) and punishment (response effort) gives the same result. As the ratio is increased, reinforcement will still tend to drive the response rate toward its upper limit. However, the cost of the food in terms of response effort increases, lowering the overall reinforcing "value" of the transaction and reducing the rate of responding below that expected if no effort were required in order to obtain the food. (There is also the effect of delayed reinforcement to be considered, which I ignore here for simplicity.)

You don't see the negative feedback regulation of food rate in the control model because food rate never reaches its reference value. For this reason I feel that your control model of the right portion of the Motheral curve does not apply in the way it would if control could be achieved.

As the reinforcement earned via lever-pressing declines with further increases in ratio requirement, one must become concerned with the possible effect of alternative sources of reinforcement on responding. The effort expended per reinforcer gets larger and larger as the ratio increases (the cost of a reinforcer), diminishing the overall reward value. There comes a point where doing other things is more rewarding than lever pressing, and behavior will shift away from the latter activity. The further diminishing return for lever pressing and availability of alternative sources of reinforcement (exploration, etc.) contribute to the left portion of Motheral's curve.

> To say that a reinforcement increases the probability of a behavior thus can legitimately be taken as meaning that R increases the probability that the _right kind_ of behavior will appear. Instead of rearing up against the wall of the cage in one place, the rat rears up in another place such that the act of rearing up causes the lever to be depressed by the front paws. What we observe is an increase in the probability that the rat will rear up in that one critical position rather than in any other. Equivalently, since we can't directly observe probabilities, we can say

that the rat's behavior changes so it shows an ever-greater proportion of rearing-up actions in the right place relative to all other places.

This is the sense in which Thorndike applied the Law of Effect to his cat data and has always been recognized as an effect of reinforcement (although this fact has sometimes been overlooked by some theorists); in addition to selection, though, reinforcement has been considered to strengthen the response absolutely as well as relatively.

> The one area of confusion that is left concerns measuring behavior in terms of rate of occurrence. This confusion shows up because in the attempt to explain performance as well as acquisition of behavior, experimenter-theoreticians set up experiments in which it was impossible, even when the right behavior had been acquired, to vary the _amount_ of effect produced by a single behavior. The use of levers and keys effectively removed any ability of an animal to have a quantitative effect on the reinforcement by varying its efforts.

This is true of most (but not all) operant conditioning experiments, where rate of response is the most typical measure of response strength. But there are many other examples in which other measures were employed, such as response force, response latency, response probability (in choice studies), and number of errors committed.

> The reinforcement model can _also_ be set up to produce a continuous relationship between rate of bar pressing and rate of reinforcement, through the intermediate step of effects of reinforcement on probability of behavior per unit time. However, this creates a conflict between the reinforcement model and the negative feedback model. The reinforcement model says that an increase in reinforcement rate must produce an increase in the probability (frequency) of behavior, while the negative feedback model says that an increase in reinforcement rate must produce a _decrease_ in the behavior rate. If this conflict can be resolved, there will be no further difficulties between the reinforcement model and the negative feedback control system model. This is not to say that neither model will be further modified on the basis of other considerations, but at least this direct contradiction will have been removed.

> The contradiction can be removed by saying that the reinforcement model applies strictly to the process of increasing the relative frequency with which the right behavior occurs, in comparison with all non-reinforcement-producing behaviors that might also occur. Once the right behavior has been acquired, the negative feedback control system model applies to the process of creating the desired amount of reinforcement.

As I hope I've communicated above, the reinforcement model can handle the data without restricting its application to the lower end of Motheral's curve. However, the application at this end does partly involve considering the selective effect of reinforcement when alternative sources of reinforcement are available for different behaviors, as you suggest.

> In the data that appear ambiguous, supporting the reinforcement model at one extreme and the control model at the other, the difference can now be explained easily, and in a way that can be tested against experimental data. Where reinforcement rates are low enough, the reinforcement model applies, and the animal begins to search for other behaviors that will more reliably produce the reinforcer. This means that other behaviors beside the bar-pressing will be seen, and proportionally less time will be spent doing the right behavior. This shows up as an apparent drop-off of behavior with decreasing reinforcement rates, or an apparent increase in behavior rate (of the kind being measured) with increasing reinforcement rates. Where reinforcement rates are high enough, the animal essentially always uses the right behavior, and controls the amount of received reinforcer near a specific reference level by varying its rate of behavior: now an increase in reinforcement rate goes with a decrease in behavior rate.

This is a nice resolution, and essentially agrees with the analysis I presented in pure reinforcement terms. Clearly the upper limb of Motheral's curve is better explained by considering the way the mechanism of the control model would be expected to operate under the conditions of the study than by the complex analysis that must be developed under reinforcement theory (although one CAN be constructed). And as you suggest, the lower limb can be understood at least partly by considering the selective property of the reinforcement principle. An interesting project would be to consider how such a selective effect might be handled by PCT.

The crux of the matter is this: Given that there are several activities that can be carried out, but which cannot be carried out simultaneously, what determines which control system will be active? For example, will the food-earning system predominate or the one that results in exploration? What determines how the organism will allocate its behavioral resources?

Regards, Bruce

Date: Fri, 23 Jun 1995 10:31:25 -0700
Subject: Re: Reinforcement theory

[From Rick Marken (950623.1030)]

Me to Bruce Abbott:

> What do you make of it?

"It" being:

```
                                                  x = random
                                                  o = reinforcement
                                                  z = subject
        |
  start |xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
        | o
     1  |  ozzzzzzzzzzzzzzzzzzzzzzz
  target|   o oooooooooooooooooooooooooooooooooooooooooooooo
y       |                          z
        |                           z
     2  |                            zzzzzzzzzzzzzzzzzzzzzzz
        |
        |
        -----------------------------------------------------
                                  Time
```

Bruce Abbott (950623.0715 EST) --

> It's a control system that doesn't control? Obviously I'm missing something here. Help me out.

The cursor stays near position 1 (now indicated on the y axis) for a while and then moves to position 2. This is a very regular phenomenon, though the transition from position 1 to 2 happens at slightly different points in the run each time; and position 2 is not always exactly the same. What I want to know is how the consequences of responses make this apparently spontaneous change in cursor position happen?

> I can't imagine a "consequence" without a change, can you?

Yes. A consequence is a result of responding; it is what happens because a response occurred. In the R. coli demo, two things can happen as a result of a response; the cursor can move up or it can move down. Whether the cursor had been moving up or down prior to the response has no effect on whether the cursor goes up or down after the response. So, whether there is an observed _change_ in the direction of cursor movement after a response is irrelevant to whether or not the direction of cursor movement after the response is a consequence of the response.

> if your cursor is drifting to left, away from target, you press the button, and nothing changes, the response has had no consequence and has been neither punished nor reinforced

(The cursor actually moves up and down, not left and right but no matter). The response has no apparent consequence (to you); but, as I said, the direction of movement IS unquestionably a consequence of the response; this is simply a FACT about the way the R. coli program relates responses to cursor movement; direction of cursor movement is a (random) consequence of responses.

> Seems to me we didn't know what the controlled variable was until we developed the models and then determined empirically which one generated a better fit to the data.

To the extent that we did this, it was a roundabout way of doing the Test. A more straightforward approach (which I _think_ I took) would have been to hypothesize a controlled variable (x+y, x/y etc) and watch for lack of effect of disturbance.

> The actual variable being controlled might be something different yet, like a weighted combination of difference and length. So don't tell me you always "know" what the controlled variable is when you develop a control model. It isn't true.

I didn't mean that we know, in the cosmic sense, what variable is REALLY controlled. But we have a VERY good idea what that variable is if we have done the Test correctly . If disturbances have very little effect on a hypothesized controlled variable (say, x+y), then that variable is a good candidate for the variable that is under control. If disturbances have even LESS effect on a different hypothesized controlled variable (say, a(x+y) + b(x/y), as you suggest), then that variable is probably a better representation of the variable that is under control.

> Yes, and there is a similar independent means of determining whether a consequence is a reinforcer, punisher, or neutral event under given conditions. Thorndike described it in 1898.

Could you explain how Thorndike would determine whether a consequence of pressing in the R. coli study was a reinforcer, punisher or neutral event? This would help a lot. Once we knew how to do this we can stop guessing about the reinforcer in the R. coli demo and determine for certain which consequences of responses are the actual reinforcers, punishers and neutral events.

Best Rick

Date: Fri, 23 Jun 1995 15:39:40 -0500
Subject: R. Coli Missing Info?

[From Bruce Abbott (950423.1535 EST)]

>Rick Marken (950623.1030) --

>>Bruce Abbott (950623.0715 EST)

>> It's a control system that doesn't control? Obviously I'm missing something here. Help me out.

> The cursor stays near position 1 (now indicated on the y axis) for a while and then moves to position 2. This is a very regular phenomenon, though the transition from position 1 to 2 happens at slightly different points in the run each time; and position 2 is not always exactly the same. What I want to know is how the consequences of responses make this apparently spontaneous change in cursor position happen?

I'd like to see the PCT model of this (Turbo Pascal, please), it ought to be interesting. There _seems_ to be some unspecified higher-level control system at work here, manipulating the reference of the cursor-control system. Or perhaps it's an homunculus at work, or signals from the soul being transmitted via the pineal gland.

If it's a higher-level control system, then the variable involved needs to be included in the reinforcement model, doesn't it? I would have to know what is reinforcing the behavior of changing the target position and thus violating the instructions given for the task (i.e., keep the cursor on the visible target as much as possible). Changing the (actual) target location is a response, too, and has its own reinforcer, not the one involved in keeping the cursor close to target.

Let's perform this experiment: Set up a compensatory tracking task and have a participant try to keep the cursor on the visible target against disturbances. You whisper in the participant's ear to change the reference location from the target to 1 cm left of the target. However, I have arranged a contingency such that the participant earns $1.00 per second so long as the cursor stays within 0.25 cm of the visible target. The participant knows this. You are offering nothing in return for compliance with your suggestion. Prediction?

Further questions: What is the reinforcer for staying on target (the visible one)? What happens if the participant adopts your suggested new target position?

I now whisper in the participant's ear. As if by magic, the cursor now stays mostly over a point about 1 cm left of target. What I have told the participant is the following: "The target region for earning the cash has now changed to where Rick suggested." What caused the change in average cursor position? Why didn't the participant change target locations when YOU suggested it?

Regards, Bruce

Date: Fri, 23 Jun 1995 16:09:38 -0600
Subject: Re: Reinforcement model and PCT model

[From Bill Powers (950623.1410 MDT)]

Bruce Abbott (950623.1135 EST) --

> It is easy to modify the model, however, so that it no longer tends to infinity. For example, think of reinforcement as being equivalent to charging a capacitor. In this case the increment in response strength is proportional to the difference between its present strength and some maximum possible strength.

Now you have to define "response strength," and give it special properties that just compensate for the instability. Better model this so I know what you're talking about. I agree that once such a problem is noticed, you can always tack on some ad-hoc feature to the model for the sole purpose of getting rid of that specific problem.

> The discriminative stimulus should not be viewed as a stimulus that elicits a response. Pr(B|S) is not the probability that S will lead to B but the probability density over time of B, given that S is present. If this probability is high and S is short, the result may look like S-R: S occurs and R follows almost immediately. But this is only an extreme case.

> ... the use of probabilities is not "dictated" at all, as one may work directly with the rate measure.

In that case you have a problem unless you add some ad-hoc feature to the model to prevent it from running away. "Running away" does not mean going to infinity; it means that there is no stable state between the physical limits of the variable.

> In cases where what varies is the magnitude of the response (i.e., response force), the reinforcement principle has been applied successfully without recourse to a probability interpretation.

How? Was this done without introducing any artificial limit or nonlinearity to prevent runaway?

> This is why reinforcement theorists talk about reinforcement affecting the "strength" of behavior rather than its probability.

But isn't response strength measured in terms of resistance to extinction? That's different from saying that a response to a stimulus exerts a force of 1 gram or 10 grams on a key. How would you put this response strength into a physical model?

> Rate or probability are considered to be measures of response strength, but so would the force exerted on the lever if the reinforcement contingency required a specific force before delivery of the reinforcer.

What would you call it if the force exerted on the lever continuously affected the amount of reinforcement delivered? Would you also have a reinforcement strength in that case? What's the physical correlate of "strength?"

> As each response continues to produce the "same" reinforcement, the rate of behavior will increase to some value dependent on the value of the reinforcer to the subject and the cost of responding (e.g., effort).

OK, so you have to make the behavior rate depend on the value of the reinforcer to the subject, minus the cost of responding. How do you introduce the value of the reinforcer to the subject in the model? Does this require adding something to the model equivalent to a reference signal?

> The "amount" of reinforcement conceived as reinforcement per response is constant in this scenario and this relationship determines how much response "strength" is incremented with each successive reinforcement, up to its limit.

If this incrementing is linear, then you still have runaway if the loop gain is greater than 1. There is still no stable state between 0 and the limit. You have to introduce just the right nonlinearity to get the response rate and reinforcement rate to level off at a finite value between the limits.

> This "molecular" analysis based on the immediate consequence of each response can be supplemented with a "molar" approach based on rates if one takes the view that the subjects in these experiments can integrate events over time. In this view, receiving individual reinforcements at a higher rate is more reinforcing than receiving them at a lower rate. A brief increase in response rate (e.g., accidental lever presses as byproducts of general activity) result in an increase in reinforcement rate (increased rates of lever- pressing are reinforced by increased rates of food pellet delivery). Decreased rates of lever-pressing lead to decreased rates of food pellet delivery (punishment). The molar approach thus predicts that response rates will tend to increase toward the maximum sustainable rate as determined by reinforcer value and response cost.

Now we're back to the imaginary case in which an increase in reinforcement rate goes with an increase in behavior rate. In general, what actually happens in most cases is the reverse of what you describe above. If an animal is feeding itself on a FR1 schedule and it accidentally hits the bar to produce extra reinforcement, it will slow its behavior to return to the previous average relationship between behavior rate and reinforcement rate (on a fixed schedule).

> This is what is observed, and not the bi-stable behavior you predict on the basis of the positive feedback relationship as you conceive it. The analysis you propose would lead to instability of behavior on a ratio schedule: any decrease in response rate (perhaps resulting from the rat's momentary need to scratch an itch) would be expected to lead to total

> extinction of the response. It doesn't happen, and it isn't predicted to happen by reinforcement theory.

You may have misread my post. I said that the instability will occur only if the loop gain (k1 * k2) is greater than 1. If it is less than 1, the effects will be amplified but there will be no runaway. When no runaway is observed, the loop gain must be less than 1.

>> To avoid this obviously inappropriate result, the meaning of "more" had to be modified so that the outcome was not a runaway condition.

> But this doesn't solve the problem, does it? Your analysis still has the instability problem I pointed out whether "more" is an increase in rate or an increase in probability.

If the loop gain is less than 1, no instability problem exists.

> Let's apply this model to the steady-state situation involved in performance on a simple CRF (1 response per reinforcement) schedule as ordinarily studied in the operant chamber. A hungry rat is placed in the chamber. We might conceive of the set point for rate of food pellet consumption as, say, 30 pellets per minute (30 ppm) under these conditions (essentially continuous eating). However, the apparatus limits the maximum rate to 10 ppm because of the delays involved in depressing the lever, moving to the food cup, picking up the food, devouring it, and returning to the lever. Thus, there is NO WAY that the rat can reach its reference level for this quantity, although it can reduce the error by a significant amount through lever-pressing (and by no other means). I believe that control theory predicts that the rat (once it has learned what to do) will develop a rate of responding that minimizes the error, up to the point where the error is reduced enough to bring the output below maximum. Depending on the gain, there will be a region within which the rate of responding will be a direct function of the magnitude of the error.

All right, you're setting up a special situation in which the rat can't reach its reference level for nutritive input because there are physical limitations, combined with the reward size, that prevent its doing so. The exact prediction that a control model would make would depend on the loop gain of the rat (measured without these restrictions), on the measured reference level for food intake, and on the degree of the restrictions. If the rat had a high loop gain, we would expect the behavior to remain at the maximum rate under all conditions.

Let's graph the situation:

```
       450*
          *
          *     x
          *
          *          x
       300* - - - - -  +
BEHAVIOR  *            |   x
          *            |
          *            |        x
 RATE     *            |                       ref level
       150*            |            x <--error--> |
          *            |                          |
          *            |                  x       |
          *            |                          |
          *            |                        x |
          *************|*************|************|**********
                       10            20           30
                    REINFORCEMENT  RATE
```

The line of x's represents the minimum consistent output sensitivity of the control system in units of presses per minute per reinforcement per minute of error. To fit your conditions it has to be about 30 (reasonable).

> If we now increase the ratio requirement, the maximum rate of reinforcement available on the schedule is less (say, 5 ppm).

It would have been easier if you had specified the new ratio requirement and then figured out the result, but we can still get there. If the new rate of reinforcement is now 5 reinforcements per minute, we can deduce that the new ratio is 390/5 = 78.

```
       450*
          *
       390* - - x
          *     |
          *     |    x
       300*     |
BEHAVIOR  *     |         x
          *     |
          *     |              x
 RATE     *     |                              ref level
       150*     |                   x <--error--> |
          *     |                                 |
          *     |                         x       |
          *     |                                 |
          *     |                               x |
          ******|********************|************|**********
                5      10             20            30
                    REINFORCEMENT  RATE
```

> This means that a larger error remains after responding has reached its maximum rate than was the case with the CRF schedule. If response rate on the CRF schedule was already at maximum, it will still be at maximum on the new schedule; if not, then it will increase.

Right.

> But this analysis ignores response cost. Assume there is a second control system for "response effort" and that the reference for this perception is set to a low value. We now have two control systems in conflict when the rate-of-eating system brings the rate of lever-pressing up and thereby brings response effort above its set point. The result is that neither system can reach set point and rate of responding stabilizes at a kind of equilibrium value between the two set points.

Actually, if you model this (as I have done), you will find that the only effect is a lowering of the effective responses-per-unit-error curve's slope. The slope does not reverse; it doesn't even become horizontal even if you make the cost very high. I was doing this in an effort to get the line of x-s to develop a curvature and go down again as the ratio requirement increased, to fit the Staddon/Motherall data. It doesn't even help to make the cost vary as the square of the response rate; the curve still never bends downward.

The only way to get the curve to bend downward is to use a cost-benefit control system as you suggest, but set its reference level so it doesn't come into effect until the error has reached about 60% of the maximum possible error. Then by adjusting the gain of this loop you can get the whole curve, over ratio requirements from 1 to 160, to match the real data very closely.

> You don't see the negative feedback regulation of food rate in the control model because food rate never reaches its reference value. For this reason I feel that your control model of the right portion of the Motheral curve does not apply in the way it would if control could be achieved.

As the above curves show, there is negative feedback regulation throughout. However, your example used unrealistic numbers. In the S/M data, the responses-per-unit-error slope was about 20. Increasing the ratio requirement from 1 to 2 decreased the reinforcement rate by about 5% (from 200 to 190) and doubled the response rate (from 200 to 380), roughly. Since the maximum observed response rate was about 3000, there is no question of being near or above the maximum response rate limit.

I would prefer that in giving examples, you would use real data. Also, when you assume the effect of some added feature of the system, it would be a good idea to set up the model and run it, to see if your intuitive prediction of what would happen actually happens.

> As I hope I've communicated above, the reinforcement model can handle the data without restricting its application to the lower end of Motheral's curve. However, the application at this end does partly involve considering the selective effect of reinforcement when alternative sources of reinforcement are available for different behaviors, as you suggest.

I think you need to actually set up the reinforcement model and run it under conditions of various ratio requirements, with and without the "alternative behaviors" provision (which applies, as you say, to the left side of the figure). In your above example, you assumed that the reversal of the relationship on the right side was due to a limitation on the maximum response speed of the organism; as a quick look at the S/M data will show you, that was not a consideration. I need to see a reinforcement model that will behave as in the right side of the S/M data as the ratio requirement is changed in the range from 1 to about 40. I already know how a control-system model will behave; for comparison, we need the behavior of a reinforcement model.

> This is a nice resolution, and essentially agrees with the analysis I presented in pure reinforcement terms. Clearly the upper limb of Motheral's curve is better explained by considering the way the mechanism of the control model would be expected to operate under the conditions of the study than by the complex analysis that must be developed under reinforcement theory (although one CAN be constructed). And as you suggest, the lower limb can be understood at least partly by considering the selective property of the reinforcement principle. An interesting project would be to consider how such a selective effect might be handled by PCT.

Good. I suggest that we make this analysis an explicit goal of the rat experiments that are just about to be done. As to the PCT analysis of the left-side data, I already have an ad-hoc PCT model of that, but let's not worry about that for now. In the paper, we can simply point out this area of comparison as being left for future study.

If a reinforcement model CAN be made to fit the right side, I suggest strongly that you go ahead and develop it, however complicated it gets. The very complication, in comparison with the simplicity of the control model, may carry a message of its own. And if it should prove that you can't find any reinforcement model that, under the observed conditions and observed range of possible behavior rates, will fit the data, that will carry an even more important message.

Best to all Bill P.

Date: Fri, 23 Jun 1995 15:37:08 -0700
Subject: Re: R. coli missing info

[From Rick Marken (950623.1530)]

Bruce Abbott (950423.1535 EST) --

> I'd like to see the PCT model of this (Turbo Pascal, please)...If it's a higher-level control system, then the variable involved needs to be included in the reinforcement model, doesn't it?

I didn't realize that reinforcement theory was based on PCT;-) What have "higher level variables" got to do with reinforcement theory? Reinforcement theory says (I think) that responses are selected by their consequences (reinforcements, punishments and non-reinforcements); the position of the cursor in the R. coli task depends on responses that are presumably selected by their consequences. How might selection of consequences produce the results we see?

> I would have to know what is reinforcing the behavior of changing the target position

That's what I have to know, too! Any ideas about where to look for what is reinforcing this behavior? I've described every possible consequence of responding in this task. There were only two possible consequences of responding; the cursor either moves away from the target after a press or toward it.

> Let's perform this experiment:...I have arranged a contingency such that the participant earns $1.00 per second so long as the cursor stays within 0.25 cm of the visible target. The participant knows this. You are offering nothing in return for compliance with your suggestion. Prediction?

From a PCT perspective, it depends on what the subject wants.

> Further questions: What is the reinforcer for staying on target (the visible one)?

I believe a reinforcement theorist _should_ say that the reinforcer is the consequence of responses that increases the probability of the responses that produce that consequence (what a reinforcement theorist actually says is quite another matter). So it would be the cursor movements (presumably movements toward the target) that increase the probability of the handle movements that produce them.

> What happens if the participant adopts your suggested new target position?

The cursor will stay near my suggested target position.

> What caused the change in average cursor position?

Depends on what theory you buy. PCT would say the change was caused by a change in the subject's reference specification for the perceived position of the cursor. Reinforcement theory would say that the change was caused by a change in the consequences of responses.

By the way, you said in an earlier post:

> there is a similar independent means of determining whether a consequence is a reinforcer, punisher, or neutral event under given conditions. Thorndike described it in 1898.

I think it is CRUCIAL that we know how this is done. How do we determine whether a consequence is a reinforcer?

I'm under the impression that we determine whether a consequence is a reinforcer, punisher or neutral event under given conditions by determining whether the consequence leads to an increase (if the consequence is a reinforcer) a decrease (if the consequence is a punisher) or no change (if the consequence is a neutral event) in the ambient probability of the response that produced the consequence.

If this is wrong, please give details (quantitative if possible) on the correct way to determine whether a consequence is a reinforcer.

Best Rick

Date: Fri, 23 Jun 1995 19:27:50 -0600
Subject: A Big Strong Rat

[From Bill Powers (950623.1850 MDT)]

Bruce Abbott (950423.1535 EST) --

I can't help horning in and seizing the opportunity.

> If it's a higher-level control system, then the variable involved needs to be included in the reinforcement model, doesn't it? I would have to know what is reinforcing the behavior of changing the target position and thus violating the instructions given for the task (i.e., keep the cursor on the visible target as much as possible).

That's the nice thing about doing away with superstitions like reinforcement. NOTHING is reinforcing the behavior of changing the target position and violating the instructions. The subject did that because that is what the subject wanted to do, to accomplish some purpose not yet announced. The instructions have no more power over the subject than the target does. It is not necessary for there to be a reinforcer to account for any behavior, or even a stimulus when it comes to that (another superstitious idea). We don't need these intervening variables to account for behavior using PCT. We simply accept the goals we find to be in effect, and the control systems that accomplish them by the means available.

> Changing the (actual) target location is a response, too, and has its own reinforcer, not the one involved in keeping the cursor close to target.

No, it's not a response and it has no reinforcer. Changing the target position is an _action_, not a response. Actions occur when reference levels change -- i.e., when intentions change. Nothing outside the organism causes that. There is no stimulus, so calling the action a response is simply a mistake. A response to what?

> Let's perform this experiment: Set up a compensatory tracking task and have a participant try to keep the cursor on the visible target against disturbances. You whisper in the participant's ear to change the reference location from the target to 1 cm left of the target. However, I have arranged a contingency such that the participant earns $1.00 per second so long as the cursor stays within 0.25 cm of the visible target. The participant knows this. You are offering nothing in return for compliance with your suggestion. Prediction?

I, the participant, will do what the whisperer suggests just to show that the reinforcement-devotee is wrong (I don't believe he'd pay me that much, anyway, unless he wanted something else from me). Or maybe I'll decide that I'm tired of this whole thing and not do what either person wants me to do. Or maybe I'll ask Rick how much he is offering. Reinforcers and contingencies don't control my behavior unless I happen to want the reinforcer at the moment, and don't find the contingency too annoying to play along with. Usually when I get tired of such games I just grab what I want of the reinforcer. I'd ask you to prove that you could really pay me all that money, and when you pulled out your wallet I'd just take the cash out of it (I'm bigger than most behaviorists). No checks accepted. If all I want out of the experiment is the money, why should I futz around with all this rigmarole you want to see going on? Gimme the money.

> Further questions: What is the reinforcer for staying on target (the visible one)?

There isn't any. That's a superstition. What makes the cursor stay on the target is that I intend for it to stay on the target, for reasons you may never know about. I might tell you for $60 an hour, payable next Tuesday or maybe next March.

> I now whisper in the participant's ear. As if by magic, the cursor now stays mostly over a point about 1 cm left of target. What I have told the participant is the following: "The target region for earning the cash has now changed to where Rick suggested." What caused the change in average

cursor position? Why didn't the participant change target locations when YOU suggested it?

Sorry, you're assuming that everything went the way you wanted it to in the first experiment. Actually I would pay no attention to your offer. I would go on tracking the target because I like Rick and wish to see him calm and happy, and I find your suggestions improper. What are you going to want for your $1.00 per second? My mother told me about people like you.

Actually, Rick offered me $10 per second, unbeknownst to you. I didn't believe he's pay me anything, either, but I did what he asked just because that's what I decided to do. It's fun just to watch you and him fight. When the experiment is over, I'm going to take both of your wallets anyway, and use the credit cards.

How do you like dealing with a BIG STRONG rat?

Best, Bill P.

Date: Sat, 24 Jun 1995 10:35:17 -0500
Subject: A Big Strong Theorist

[From Bruce Abbott (950624.1030 EST)]

>Bill Powers (950623.1850 MDT) --
>>Bruce Abbott (950423.1535 EST)

Now that I've recovered from paroxysms of laughter, I believe I'm ready to respond to your post concerning "a big strong rat." So, down to business...

> I can't help horning in and seizing the opportunity.

I'm not surprised--but let's not get confused about what we're doing here. What I was trying to do was to indicate how reinforcement theory would handle the effect described. I know how PCT does it. This is not an argument about which view is correct, but rather, it is an attempt on my part to impart some knowledge about reinforcement theory. I believe a proper understanding of the theory is essential if it to be successfully attacked.

>> If it's a higher-level control system, then the variable involved needs to be included in the reinforcement model, doesn't it? I would have to know what is reinforcing the behavior of changing the target position and thus violating the instructions given for the task (i.e., keep the cursor on the visible target as much as possible).

> That's the nice thing about doing away with superstitions like reinforcement. NOTHING is reinforcing the behavior of changing the target position and violating the instructions. The subject did that because that is what the subject wanted to do, to accomplish some purpose not yet announced. The instructions have no more power over the subject than the target does. It is not necessary for there to be a reinforcer to account for any behavior, or even a stimulus when it comes to that (another superstitious idea). We don't need these intervening variables to account for behavior using PCT. We simply accept the goals we find to be in effect, and the control systems that accomplish them by the means available.

I understand that this is your position, which is derived from PCT. But it is also your position that reinforcement theory cannot apply here, and this is incorrect. Let's examine your statements carefully:

> That's the nice thing about doing away with superstitions like reinforcement. NOTHING is reinforcing the behavior of changing the target position and violating the instructions. The subject did that because that is what the subject wanted to do, to accomplish some purpose not yet announced.

From the point of view of reinforcement theory, SOMETHING must be producing this behavior; the question is what. The answer, for the radical behaviorist, is to be found in the current environment and in the lasting effects of the organism's history of experience with the environment and its contingencies. If the participant "wants" to keep the cursor on the visible target, it is because this history and the contingencies currently in effect (as perceived by the participant as a result if his or her experience) indicate that such behavior will be reinforced under the present circumstances. So, in theory at least, If I could become privy to the appropriate details of your personal history and the sources of reinforcement in the present environment, I could fully account for your failure to keep the cursor on target as instructed. In practice, I am not in a position to know all the relevant details, although I can provide an educated guess about some of them based on the limited evidence available.

> The instructions have no more power over the subject than the target does. It is not necessary for there to be a reinforcer to account for any behavior, or even a stimulus when it comes to that (another superstitious idea). We don't need these intervening variables to account for behavior using PCT. We simply accept the goals we find to be in effect, and the control systems that accomplish them by the means available.

Reinforcement theory, also, would "simply accept the goals we find to be in effect" given our ignorance as to the conditions that brought them about. This is not to say that we would be satisfied with guessing about them, not if our aim is to provide a scientific explanation as opposed to merely providing a possible account that fits the data. We would devise experiments to identify the source or sources of reinforcement, the reinforcement equivalent of performing the Test in PCT. However, my purpose here is not to prove that your behavior is controlled by some particular source of reinforcement, but to show that reinforcement theory can offer a reasonable, internally consistent explanation for that behavior.

In searching out these "sources of reinforcement," reinforcement theorists are at least taking the position that there are causes of the behavior that can be identified. In your analysis based on PCT, you seem to be saying that the organism does what it does because it "wants to," without suggesting why.

>> Changing the (actual) target location is a response, too, and has its own reinforcer, not the one involved in keeping the cursor close to target.

> No, it's not a response and it has no reinforcer. Changing the target position is an _action_, not a response. Actions occur when reference levels change -- i.e., when intentions change. Nothing outside the organism causes that. There is no stimulus, so calling the action a response is simply a mistake. A response to what?

You are giving me the PCT view again, rather than attempting to understand the reinforcement view.

Reinforcement theorists recognized long ago that certain processes going on in the brain could not be observed from the outside, e.g., making a decision. One solution is to ignore these processes and refer instead strictly to the organism's history of reinforcement, on the assumption that these unobservable intermediaries must be the products of such a history and therefore that the observable relationships can be accounted for just as well by staying with the external observables (the Skinnerian approach). Another solution is to assume that these internal processes work just like the external, observable ones do. Why invoke new processes if you can assume that the same processes are at work internally as externally? This solution has the advantage of parsimony, although testing for such internal processes is necessarily indirect.

> No, it's not a response and it has no reinforcer.

I can certainly observe that you're keeping the cursor somewhere other than the place where you were instructed to keep it. That's a response. As to the reinforcer, what about this one:

> I, the participant, will do what the whisperer suggests just to show that the reinforcement-devotee is wrong (I don't believe he'd pay me that much, anyway, unless he wanted something else from me).

So, because of your personal reinforcement history, you (a) find showing reinforcement-devotees wrong to be reinforcing and (b) do not perceive a contingency between staying on the experimenter's selected target and earning the cash. It is entirely understandable, from a reinforcement theory point of view, why your cursor does not stay on the designated target.

> Actions occur when reference levels change -- i.e., when intentions change. Nothing outside the organism causes that.

And reference levels change because.....? I intend to mail a package at the post office this afternoon. On arrival, I discover that the post office is closed. So I give up my intention, and intend to go home instead. Nothing outside me has caused that? Hmmmm....

> There is no stimulus, so calling the action a response is simply a mistake. A response to what?

Operants are simply emitted; they require no stimulus, not since 1938. Call them operants rather than responses if you like, it's all the same to me.

In the remainder of this post, you offer a series of possible reasons why you would not conform to the experimenter's wishes, even when given a relatively strong inducement to do so ($1.00 per minute). You seem to think that this behavior shows that reinforcement theory is wrong. It's a common criticism of reinforcement theory in applied settings that reinforcement "doesn't work" because Johnny still misbehaves in class even though the teacher offers Johnny a bribe for his compliance. The problem, from the point of view of reinforcement theory, is not that reinforcement doesn't work, but that there are other, more powerful sources of reinforcement at work which are controlling Johnny's behavior, such as the attention he gets from misbehaving. You've named several of those sources for your misbehavior:

> Actually I would pay no attention to your offer. I would go on tracking the target because I like Rick and wish to see him calm and happy, and I find your suggestions improper. What are you going to want for your $1.00 per second? My mother told me about people like you.

> Actually, Rick offered me $10 per second, unbeknownst to you. I didn't believe he's pay me anything, either, but I did what he asked just because that's what I decided to do. It's fun just to watch you and him fight. When the experiment is over, I'm going to take both of your wallets anyway, and use the credit cards.

So, as a behavior analyst, my problem is that my paltry $1.00 per minute is no match, in terms of reinforcing power, for the reinforcement you receive from Rick's approval, from frustrating people whose suggestions you deem improper (because of a reinforcement history involving your mother and her stories about "people like me"), from the money Rick offered you not to conform to my instructions, and from the fun you get when Rick and I fight about why you didn't behave as expected. To get you to do as I wish, I must attempt to (a) eliminate these competing sources of reinforcement, and (b) identify a source of reinforcement that will work in your case.

Fortunately for me, you have been kind enough to list these competing sources, which makes my job easier. In the process, you have inadvertently provided all the evidence any reinforcement theorist could want that your nonconforming behavior is indeed under the control of a number of potent reinforcers. Thanks!

> How do you like dealing with a BIG STRONG rat?

Little rats, big rats, tigers, elephants, blue whales, size makes no difference, not when you've got a BIG STRONG theorist who knows how to handle them. Pretty soon, they're all eating out of his hand . . . (;->

Regards, Bruce

Date: Sat, 24 Jun 1995 15:34:54 -0600
Subject: Reinforcement theory and PCT

[From Bill Powers (950624.1005 MDT)]

Bruce Abbott (950624.1030 EST) --

> From the point of view of reinforcement theory, SOMETHING must be producing this behavior; the question is what. The answer, for the radical behaviorist, is to be found in the current environment and in the lasting effects of the organism's history of experience with the environment and its contingencies.

With these assumptions, there is nothing left to do but to try to identify the current environmental causes and the history of interactions. The behaviorists do not really consider that the causes of behavior may be internal, and trace back to the environment only via the long path through evolution. They have renounced that option.

This insistence on observable causes suggests a weak grasp of how complex systems work. Even in a computer, it is impossible to assign specific causes to most outputs. The causes arise from interactions among system components, programs and subprograms, and programs and users; what a user enters at the keyboard may have only the most general and fuzzy relationship to what comes out at the printer. Many kinds of outputs, such as a periodic output of the time of day, may result entirely from the internal structure of the system.

The PCT and HPCT models do not try to account for behaviors in terms of causes and effects. A "want" (a reference signal) may determine the level to which a perception is brought, but the action required to change the perception as required always depends on the nature of the external feedback connection and any disturbances that happen to be acting. And the want itself is simply an output from a higher system that is using many lower ones, via manipulations of reference signals, to control a higher-order perception composed of both controlled and uncontrolled perceptions of lower order. The entire hierarchy of wants is itself a product of some kind of reorganizing system that controls variables critical to survival in relation to genetically-given reference levels. The basis "causes" of behavior can't be traced to the environment that exists in the lifetime of a single organism.

As you say, "From the point of view of reinforcement theory, SOMETHING must be producing this behavior; the question is what." The "what" answers that PCT offers are totally different from those that the behaviorist wants: the answer is that THE ORGANISM produces behavior. The environment is only the means by which the brain controls what the environment does to the organism.

This is such a fundamental difference that there is little point in any direct attempt to pit one view against the other.

If an inexplicable behavior occurs, the behaviorist does not say "Well, the cause must lie inside the organism, since I can't find it in the environment." That would contradict the basic belief that all causes are in the environment. Instead, the behaviorist says "There must have been some reinforcer acting that I didn't see." Thus both successes and failures of prediction are taken into account, and the theory is invulnerable against disproof. For instance:

> It's a common criticism of reinforcement theory in applied settings that reinforcement "doesn't work" because Johnny still misbehaves in class even though the teacher offers Johnny a bribe for his compliance. The problem, from the point of view of reinforcement theory, is not that reinforcement doesn't work, but that there are other, more powerful sources of reinforcement at work which are controlling Johnny's behavior, such as the attention he gets from misbehaving.

In a real science, when a prediction fails you don't excuse it by saying that something unobserved must have happened in just the right way to keep the theory true. You investigate why it failed and produce the data that explains the failure. The "more powerful sources of reinforcement" kind of argument can make ANY theory true no matter how often it fails. The fact is that reinforcement theory, when applied in the classroom or elsewhere, often DOES fail. That is what the _real_ data say. Invoking imaginary data to make the failure appear like a success takes you out of science and into faith. It's like explaining why, if there is a God, the world is so full of suffering babies. The answer is that God has purposes too mysterious for mere mortals to understand, and all is going according to God's plan despite how it looks to us. The behaviorist says that reinforcement works even if it appears not to be working.

What I was doing in my "600-pound rat" post was talking like a behaviorist. The method is simply to assert the correctness of one's interpretation and force it onto the other person. Instead of laying out my detailed reasons, I just said

>> No, it's not a response and it has no reinforcer.

And you replied in kind:

> I can certainly observe that you're keeping the cursor somewhere other than the place where you were instructed to keep it. That's a response... So, because of your personal reinforcement history, you (a) find showing reinforcement-devotees wrong to be reinforcing and (b) do not perceive a contingency between staying on the experimenter's selected target and earning the cash. It is entirely understandable, from a reinforcement theory point of view, why your cursor does not stay on the designated target.

This kind of battle of assertions gets nobody anywhere. If this were all there were to the dispute, we could solve it by talking like behaviorists on Mondays, Wednesdays, and Fridays, and like PCTers on Tuesdays, Thursdays, and Saturdays. On Sundays we could say it's all God's Will.

The basic question is whether behavior really works as each side says it works. Resolving that question requires thinking up tests that could give preference to one theory over the other. It requires taking failures seriously and using only observed data to explain them. If Johnny doesn't respond to operant conditioning because of other more powerful reinforcers, show what the other reinforcers were and prove that they are more powerful; don't just say they must have existed. If Johnny tries to control and fails, show what the detailed causes of the failure were, and _prove_ that they will always cause such a failure.

> So, as a behavior analyst, my problem is that my paltry $1.00 per minute is no match, in terms of reinforcing power, for the reinforcement you receive from Rick's approval,

Pay closer attention. I said that I didn't choose to do it Rick's way because of the reward, but for a different reason all my own. Of course as a behaviorist you had to assume that the larger reward prevailed, even if it didn't, and that my "reasons" were illusions.

>> How do you like dealing with a BIG STRONG rat?

> Little rats, big rats, tigers, elephants, blue whales, size makes no difference, not when you've got a BIG STRONG theorist who knows how to handle them.

This is one thing that behaviorists uniformly overlook. Size and strength make all the difference. You can apply operant conditioning to rats, children, prisoners, and inmates of mental institutions because you have overwhelming physical force at your disposal. Your subjects have no way to prevent you from establishing any contingency you choose. They have no independent source of the reinforcers you supply. They have no way to escape from the situation altogether.

You tell me you can pay me $1.00 per second, but you will do so only if I perform as you want to see me perform. If I were a child or someone cowed by your authority and surroundings and legal backing, and wanted to get that money more than I wanted not to produce this meaningless behavior, I would do what was required. But as a 600-pound rat who wanted the money and didn't give a **** about your position, all I had to do was take it from you, once I knew you had it. It's very hard to reward me for something without revealing that you have a supply of the reward and are doling it out contingent on my following your wishes. Also, it's very hard to reward me when you can't keep control over the source of the reward. Even the gerbils I once "trained" figured that out: they were climbing all over the apparatus looking for a way to get to where the pellets came from. Only my superior size and strength kept them from doing it. And they got it all anyway, by pressing day and night until they had emptied the hopper and stored the food away under their nest.

The whole reinforcement scenario is a game played under unspoken rules that make it work. The main rule is that the biggest guy gets to establish the contingencies for the littlest guy. The next main rule is that the biggest guy has control of the source of reinforcers, and the littlest guy can't get it away from him.

Best, Bill P.

Date: Sat, 24 Jun 1995 18:41:55 -0600
Subject: Task for comparing PCT and reinforcement theory

[From Bill Powers (950624.0620 MDT)]

Bruce Abbott and Rick Marken --

It occurs to me that we've been missing our best bet for a task to compare PCT and reinforcement theory: a simple compensatory tracking task. Let's set up a somewhat novel version in which cursor position is set by

C := C + k*(H + D)

We know that with a nice slow disturbance, the subject will be able to keep the cursor very close to, although not exactly on, a stationary target. The errors, however, will show a very low correlation with the handle movements. We can use a running average of error-squared, shown as 100 minus the average error, as the "reinforcer." The cursor position relative to the target, presumably, would be the discriminative stimulus.

We know what the general PCT prediction is: handle position equal and opposite to disturbance, cursor remains near target. The reinforcement explanation would have to be something like probability of rightward handle movement given cursor left or probability of leftward handle movement given cursor right increases when the score moves toward 100, decreases when it moves away from 100. Bruce can work out the details.

What will make the difference is that with an easy disturbance the person will be controlling with an error near the noise level. The PCT model will predict control whether the reinforcer is present or not -- or even if it varies randomly (lying about the actual error). An actual working model using reinforcement theory will control very poorly because of the low correlations, even with a truthful reinforcer, while the PCT model will control as well as the subject does. I think.

This is a velocity-control task, so the PCT model just needs a proportional gain in the output function; a very simple model with one parameter will do well enough.

If this task proves too easy, we can add another derivative:

dc := dc + k*(H + D)*dt

c := c + dc*dt

This will require a first derivative in the output function of the control system.

Best, Bill P.

Date: Sun, 25 Jun 1995 12:00:31 -0400
Subject: Re: Reinforcement Theory

<[Bill Leach 950625.10:44]
>[From Bruce Abbott (950623.0715 EST)]

I REALLY don't want to "get between you and Rick"... since the EPA banned asbestos, and it's quite hot around here already...

> I can't imagine a "consequence" without a change, can you? If the ...

I realize that this is bordering of "word games" but yes, I can. If press the key and nothing happens that is a "consequence". The "consequence" is that I did something and "nothing" happened.

You may view Rick's logic for the program as faulty and possibly you are right. However, I don't think that it is "strange" to consider that a purposeful being that has "set out to do something":

explicitly performs an action and observes that the action fails to produce desired results

would be expected to conclude that the consequences of the action was unfavorable.

In one sense, I do agree with you that "consequence" can not exist without change. Where I disagree is that I believe that it is illogical to presume that only a change in _result_ is a consequence. The "initiating action" is also a _change_.

If I push on a door and it does not open, you are claiming that there were no consequences to my action -- that is preposterous!

> advanced knowledge of the CEV

You "hooked" him pretty good on that one Bruce!

EAB does acknowledge that "reinforcement", "punishment" and "neutral events" are specific to the individual experimental subject under the specific conditions of the experiment, yes? That is, EAB researchers recognize that they can not accurately "generalize" the experimental results?

For example, one of the concepts of EAB (that I might well be rather confused about), is the idea that a "reinforcer" causes repetition of the behavior that "resulted" in the reinforcer.

Nearly anyone more than 8 years old would see that the above is just plain not true (including EAB researchers of course).

So there is a concept of "satiation"(?) added to account for some of the observed limits for effect of reinforcers upon behavior. Is not "satiation" an outright admission that the "reinforcer" is only related in some actually unknown way to whatever process is being observed? Is not satiation admitted to be "somehow" internally controlled by the subject? Specifically, does not EAB recognize that their own data proves that "reinforcers" can not be the cause of behavior?

It is true, is it not, that in an "action get you a food pellet" type experiment that the delivery of four pellets vice one pellet per action results in fewer actions? What happens if each action produces a reinforcer that is greater than that amount that the subject would eat for the duration of the experiment?

It seems to me that such presents a significant challenge to the basic principle of EAB. Every "modification" points to the idea that the external "thing" IS NOT determining the subject's actions but rather what the subject wants determines the subject's actions. Thus, the "reinforcer" is incidental - if the subject wants some amount of the reinforcer then behavior will adjust (if possible) to get that just that amount.

The matter is so profoundly different when compared to PCT research as to be incomparable (even when it appears otherwise). PCT asserts that with the single exception of application of overwhelming force that the ONLY thing that is important is the wants, goals, or references of the subject. Any research that does not explicitly search for, test for and describe the postulated reference(s), obtains data that at best, can NOT be reliably generalized and at worst, is completely wrong as understood.

-bill

Date: Sun, 25 Jun 1995 12:15:30 -0500
Subject: <No subject given>

[From Bruce Abbott (950625.1210 EST)]

Well guys, today I'm starting the second half of my first century. I don't know how this has happened so soon; since when does 50 come right after 35?

Now, let's see if the remaining brain cells still work . . .

>Bill Powers (950624.1005 MDT) --

>>Bruce Abbott (950624.1030 EST)

> From the point of view of reinforcement theory, SOMETHING must be producing this behavior; the question is what. The answer, for the radical behaviorist, is to be found in the current environment and in the lasting effects of the organism's history of experience with the environment and its contingencies.

> With these assumptions, there is nothing left to do but to try to identify the current environmental causes and the history of interactions. The behaviorists do not really consider that the causes of behavior may be internal, and trace back to the environment only via the long path through evolution. They have renounced that option.

> . . .

> If an inexplicable behavior occurs, the behaviorist does not say "Well, the cause must lie inside the organism, since I can't find it in the environment." That would contradict the basic belief that all causes are in the environment. Instead, the behaviorist says "There must have been some reinforcer acting that I didn't see." Thus both successes and failures of prediction are taken into account, and the theory is invulnerable against disproof.

Let's not confuse glib "explanations" with offerings of proof. If the theory assumes that the causes of behavior are to be found in the environment (both present and past), then that is where you look for them. PCT's approach is not different: if the theory assumes that all action results from error in a controlled perception, then that is what you look for to explain the behavior. If either theory founded its "proof" merely on the ability of its practitioners to identify reasonable-sounding sources of behavioral motivation, then both would be invulnerable against disproof. When asked to speculate as to why someone is doing such-and-such, practitioners in either camp can point to variables which may be as work and would permit the theory to "explain" the behavior.

However, this is not research. When scientific research is being conducted in order to explain a given set of observations relating to behavior, considerable effort is expended to test alternative hypotheses and, to the extent possible, rule out those that fail the tests. This is done in PCT, and it is done in behavior analysis.

>> It's a common criticism of reinforcement theory in applied settings that reinforcement "doesn't work" because Johnny still misbehaves in class even though the teacher offers Johnny a bribe for his compliance. The problem, from the point of view of reinforcement theory, is not that reinforcement doesn't work, but that there are other, more powerful sources of reinforcement at work which are controlling Johnny's behavior, such as the attention he gets from misbehaving.

> In a real science, when a prediction fails you don't excuse it by saying that something unobserved must have happened in just the right way to keep the theory true. You investigate why it failed and produce the data that explains the failure. The "more powerful sources of reinforcement" kind of argument can make ANY theory true no matter how often it fails. The fact is that reinforcement theory, when applied in the classroom or elsewhere, often DOES fail. That is what the _real_ data say.

In a real science, when a prediction fails you develop tests to determine what is going on. For example, your observations may suggest that Johnny is misbehaving because of the attention such behavior brings. If this hypothesis is correct, then eliminating this source of reinforcement should allow this behavior to extinguish. So, you implement a test by introducing "time out" for misbehavior. Each time Johnny begins to disrupt the class, he is immediately taken to a very boring room and has to sit there for a few minutes watching Ms. Blatherstone type reports. If, after a few such experiences, we observe that the frequency of Johnny's disruptive behavior takes a nosedive, we have empirical support for the hypothesis. Well run behavior mod programs don't offer excuses for the failure of the manipulations to have their intended effects, they use systematic data collection to identify the source of the problem and find a way to correct it.

> The basic question is whether behavior really works as each side says it works. Resolving that question requires thinking up tests that could give preference to one theory over the other. It requires taking failures seriously and using only observed data to explain them. If Johnny doesn't respond to operant conditioning because of other more powerful reinforcers, show what the other reinforcers were and prove that they are more powerful; don't just say they must have existed. If Johnny tries to control and fails, show what the detailed causes of the failure were, and _prove_ that they will always cause such a failure.

Yes, I agree, and so would any competent behavior analyst. As I noted in my post, my pointing to these putative sources of reinforcement was done only to show that such an accounting is _possible_, not that these sources are in fact the correct explanation for your behavior. Finding the correct explanation would take research.

In fact one of the problems for both _applied_ behavior analysis and _applied_ PCT is that you don't know all the factors at work in a given individual, whether the relevant factors are thought to be the history of reinforcement and current contingencies or the relevant set of controlled perceptions up and down the hierarchy. This means that the job involves a lot of educated guesswork, based on observation and perhaps a bit of experimentation. Even then it's not always going to work, because information is incomplete and your ability to exert influence is limited.

In a basic research setting, we often can simplify the situation enormously. For example, by using naive rats, I can be fairly well assured that my subjects will not be motivated by a desire to frustrate the experimenter, as that bigger rat you mentioned appears to be, and I can do a pretty good job of eliminating most alternative sources of reinforcement, other than the one or two whose effects I wish to study.

>> So, as a behavior analyst, my problem is that my paltry $1.00 per minute is no match, in terms of reinforcing power, for the reinforcement you receive from Rick's approval,

> Pay closer attention. I said that I didn't choose to do it Rick's way because of the reward, but for a different reason all my own. Of course as a behaviorist you had to assume that the larger reward prevailed, even if it didn't, and that my "reasons" were illusions.

I did pay close attention--did you? I didn't say you choose to do it Rick's way because of the reward, I offered several possibilities, including that you find it rewarding to frustrate reinforcement theorists. As a reinforcement theorist, I would not assume that your reasons are illusions (they may or may not be), but that, whatever they may actually be, those reasons can be traced back to your experiences.

Regards, Bruce

Date: Sun, 25 Jun 1995 23:39:20 -0400
Subject: Re: Reinforcement theory and PCT

<[Bill Leach 950625.13:07]
>[From Bill Powers (950624.1005 MDT)]

I am beginning to think that the "argument" is futile.

Within limits, it is not unreasonable to conclude that the "environment makes us "behave" the way that we do behave."

That some environmental factors have a consistent, testable cause effect relationship to living beings is beyond dispute (at least with anyone even pretending a scientific approach). For example, placing a human (or any mammal) in a normal atmosphere except that it is oxygen free for a sufficient length of time will result in the organism's death. The "sufficient time" varies a bit but even some of the variation "causes" are well known.

That I use (abuse?) the english language can be said to be "caused" by my environment though with much less certainty than the oxygen relationship (also I believe, that such a statement carries great deal less significance).

From Bruce's last posting on this subject:

[Bruce Abbott (950624.1030 EST)]

I conclude that where analysis and testing result in a match between observed behavior and the predicted "cause" EAB will always (also) be correct.

EAB's "reinforcer" is PCTs CCEV. [Martin corrected me on my introduction of what possibly amounted to a high level of confusion. I was using CEV to mean "CONTROLLED Environmental Variable" as opposed to "COMPLEX Environmental Variable" and thus should have been using CCEV in most if not all places where I had used CEV.]

If a reinforcer that has been shown to be a reinforcer in the past does not "reinforce" then "there must be "bigger reinforcers that are interfering". In PCT if a CCEV, again previously shown to be a CCEV, turns out to drift or change as a result of disturbance then either the CCEV isn't or the disturbance exceeds the output capacity of the organism.

A search for "bigger" reinforcers or for the CCEV could look very much alike.

The major irreconcilable (I think) differences are:

EAB considers that learning what the reinforcers themselves are is of primary research importance in a given experiment.

EAB then concludes that if you knew the history of reinforcer experience you could predict an organism's behavior in the presence of current reinforcers.

EAB then takes the position that events in the environment are not only the ultimate cause of behavior overall but are (somewhat vaguely) the cause of current behavior.

PCT considers that it is necessary to determine the CCEV before any behavioral research can be conducted with a specific subject.

PCT concludes that you can not predict future behavior of an individual subject based upon knowledge of a history of CCEVs. You can only conclude that if a reference exists for a particular CEV then that organism's behavior will be an attempt to bring the perception of that CEV into a satisfactory match to the reference. PCT recognizes that repeated tests (same individual or different) CAN result in differences in control method for a variety of reasons. The first and possibly most obvious would be due to "learning". A second could be due to the existence of other references. Other references may or may not impact "quality of control" and may or may not impact control method but could affect either or both.

Thus, PCT makes NO prediction as to the specific details about how the organism _will_ attempt to achieve control in a particular instance. Indeed currently PCT IS the attempt to study the observed control methods.

That HPCT systems can balance conflicting (though not mutually exclusive) CCEVs virtually guarantees that method of control will not always be the same for all individuals or all tests.

PCT then takes the position that events in the environment are only the ultimate cause of behavior overall because the subject would not even exist in the environment if some events had not occurred.

The environment is taken to influence or effect behavior (but not control).

If I have a reference for "picking up mount Shasta with my hands", the "environment" will "stop me" from doing such. To PCT the "consequences" of my attempt are neither a "reinforcer" nor a "punisher". PCT does recognize that somehow my "control failure experience" is "learned". PCT does NOT predict that I will immediately discontinue attempting to pick up mountains as a result of this failure or any other. THE MERE FACT OF THE EXISTENCE OF A CONTROL FAILURE IS NOT A GUARANTEE OF BEHAVIORAL CHANGE!

Unless the reference changes or reorganization becomes involved I could happily go around for the rest of my life trying to pick up mountains blissfully unaware of "all the punishment that I am receiving". [and I recognize that I a using an assertion that is not necessarily "orthodox" PCT in that I am maintaining that a control error may not necessarily initiate reorganization]

Of course the EAB answer is that I receive a greater "reinforcement" for trying than I do "punishment" for failing - no resolution possible. In a sense EAB would be correct in this case if only they could be brought to the point where they recognize that the "reinforcement" is internal and it is ALWAYS internal.

-bill

Date: Sun, 25 Jun 1995 23:39:58 -0400
Subject: WHY?

<[Bill Leach 950625.14:50]
>[From Bruce Abbott (950624.1030 EST)]

> ... performing the Test in PCT. However, my purpose here is not to prove that your behavior is controlled by some particular source of reinforcement, but to show that reinforcement theory can offer a reasonable, internally consistent explanation for that behavior.

> In searching out these "sources of reinforcement," reinforcement theorists are at least taking the position that there are causes of the behavior that can be identified. In your analysis based on PCT, you seem to be saying that the organism does what it does because it "wants to," without suggesting why.

These two paragraphs are probably worthy of a doctoral dissertation!

"Reasonable" is certainly up to question but most debates in that area ultimately are not too useful.

"Internally consistent" is a bit more "meaty". That EAB is "internally consistent" in the sort of discussions that we have been having is rather a bit far fetched. To assert that past "reinforcers" can explain all deviations from expected results is about as scientific as astrology. Such a claim makes EAB hypothesis at best until the data is in demonstrating the truth of the claim.

Of course HPCT is not significantly better off in that regard. Both lend themselves to "logical thought experiments". I am not sure that either can really be demonstrated in an irrefutable fashion to be superior.

> The WHY

As far as proof is concerned, I don't think that either hypothesis is in a position to provide any.

PCT's only justifiable claim is to reason and observational experience. I believe that most individual's have had at least some experience personally with something that I choose to call a "religious experience". I use that terms only because the "sensations" and "attitude changes" described by people claiming to have had a "religious conversion experience" seem to "match" rather well. My own belief is that this same "feelings/attitude changes" can occur with regard to many experiences besides religious ones.

People can and do make a "fundamental shift" in their entire manner of relating to their environment. How does EAB explain a "Mother Teresa"(sp?) or better still a "Joan of Arc"? What external "thing" can have the power to cause one the give up one's own life?

For that matter, what possible characteristic of something in the environment can "make" an organism do anything unless the organism itself wants or does not want some amount of the external thing? Is EAB claiming that the external thing creates the degree of want itself?

Also, you say:

> Reinforcement theory, also, would "simply accept the goals we find to be in effect" given our ignorance as to the conditions that brought them about.

which itself is laudable enough sounding but what does that mean to an EABer? To "simply accept the goals" would seem to imply acceptance of "purposeful behavior", that is the idea that the organism does what it does for its' own reasons. That would seem to move the "power" of the reinforcer back inside again.

By "accepting" and then ignoring the goals, is not one then ignoring what one knows through their own personal experience to (at least occasionally) have an overwhelming impact on what one does and why?

> ... Why invoke new processes if you can assume that the same processes are at work internally as externally? ...

But can you assume? Is not this yet another one of those "things" that is obviously false based upon one's own experience? If my goal is to drink a glass of milk, there are quite literally thousands of different ways in which that goal might be satisfied including the many different actions that would occur should there not be any milk in the house or what exists is spoiled, etc.

The milk is not a "reinforcer". If anything is, it would be my desire but my behavior with regard to the "reinforcer" STOPS after I drink the milk!

> And reference levels change because.....? I intend to mail a package at the post office this afternoon. On arrival, I discover that the post office is closed. So I give up my intention, and intend to go home instead. Nothing outside me has caused that? Hmmmm....

You intended to have a packaged mailed. You decided to take it to the post office yourself (new reference set to implement original reference - which itself has at least a couple of higher goals).

Did you choose to throw the package away? Did you decide not to mail it at all? Post office closed is a disturbance. This particular disturbance in combination with other perceptions (such as the likely consequences of breaking in) is perceived as not directly controllable.

Our perceived reality interferes with "our plans" all the time. That we are control systems is the reason that we normally accomplish our goals anyway.

> Operants are simply emitted; they require no stimulus, not since 1938. Call them operants rather than responses if you like, it's all the same to me.

HOW are they emitted? Not why, but how? How does an organism "emit" just the exact behavior needed for a particular situation?

> .. To get you to do as I wish, I must attempt to (a) eliminate these competing sources of reinforcement, and (b) identify a source of reinforcement that will work in your case.

This is the "everybody has their price" theory that Frankl smashed flat in the fifties. Though far from common it is possible that Bill P. would even starve to death rather than "emit" the "proper" behavior.

-bill

Date: Mon, 26 Jun 1995 09:05:53 -0600
Subject: PCT and EAB

[From Bill Powers (950626.0630 MDT)]

Bill Leach (950625.13:07 U.S. Eastern Time Zone) --

You bring up a number of interesting points about reinforcement theory and PCT. One recurrent theme is that the environment does have some controlling effects on our behavior -- but then again, it doesn't!

> Within limits, it is not unreasonable to conclude that the "environment makes us "behave" the way that we do behave."

> That some environmental factors have a consistent, testable cause effect relationship to living beings is beyond dispute (at least with anyone even pretending a scientific approach). For example, placing a human (or any mammal) in a normal atmosphere except that it is oxygen free for a sufficient length of time will result in the organism's death.

But in another post:

> If my goal is to drink a glass of milk, there are quite literally thousands of different ways in which that goal might be satisfied including the many different actions that would occur should there not be any milk in the house or what exists is spoiled, etc.

One of the main points of the PCT model is to explain how it is that there are quite literally thousands of different ways in which a given goal might be satisfied, yet somehow an organism finds just the way that will do the trick -- most of the time. This goes clear back to the basic puzzle about behavior that William James noticed: variable behavior producing consistent ends.

Probably the greatest hole in reinforcement theory was patched by the most casual plastering job. Skinner noticed way back in the beginning that what we casually describe as "the same behavior as before" is most often a very different behavior from the one we saw before. The actions of organisms, in the sense of the outputs they produce with muscles and limbs, do not actually repeat from one instance of behavior to another. What repeats is some particular consequence of acting. Rather than asking how this can possible be, Skinner simply defined "the operant" as any class of actions that has a particular effect.

At that point a serious question for a materialist explanation of behavior arises: how is it that the same effect can be produced by different actions? If Skinner had been inclined to take a physical-science approach to behavior, this would have been a serious question indeed. And a physical-science answer would have been of little comfort to a behaviorist. The only way in which a consistent result can come from variable actions is for the environment to vary in just the way needed to make up for the variations in the actions.

That, of course, is putting it the wrong way around. The right way starts by asking what the same consequence would have looked like without the organism's actions. Where would the milk that you drank have been if you had not acted? It might have been in a thermos jug, in a carton in the refrigerator, in a grocery bag waiting to be unpacked, in a store, or in a cow. Yet no matter where it is, you end up getting a drink of milk. The only POSSIBLE explanation is that under each different environmental circumstance, you produced just the physical action required for the milk to end up in a glass and the glass to be emptied down your gullet. Your action does not produce the final effect. All it does is make up the difference between a particular final effect and the effect that would have happened without the action. If someone goes and gets the milk and puts it into a glass and tips the glass into your mouth, the only action required from you to get a drink of milk is to open your mouth and swallow.

The subject matter of PCT lies in the very part of behavior that Skinner dismissed as basically unexplainable. When he said that behavior is "emitted," he ceased to talk about the physical actions of the organism and began talking about _physical consequences_ of those actions. What is "emitted" is a change in the physical observable world due to a change in the organism's action coupled with any change in the physical world capable of altering that same physical variable. What we commonly call behavior is really a resultant; the outcome of combining forces created by an organism with forces that originate elsewhere.

When we see it this way, we realize how strange it is that an organism can actually appear to emit the same physical consequence of action over and over. We can't just say, as Skinner did, "Oh, well, the organism just did one of the many things that could have had the same physical effect." It's not as simple as that. In any one physical situation, there is only ONE action the organism can take that will have a particular physical effect. If the local environment changes in any way, there is still only ONE action the organism can take to create the same effect as before, but now it is a DIFFERENT action.

What's really going on is that when the organism emits a particular physical effect, its action is PRECISELY THE ONLY ACTION THAT COULD HAVE PRODUCED THAT EFFECT AT THAT TIME. Given the same state of the environment, any other action would have resulted in a DIFFERENT effect. As Skinner noted, there are many actions that can have the same effect. But what he failed to notice is that this is true only over a large number of instances of the actions. In any one instance, it is not true that any number of actions can have the same effect. Given the current effects of the environment on a particular variable, for that variable to be in or remain in a particular state requires ONE AND ONLY ONE ACTION by the organism. It is that action which, when added to all other influences acting at the same time, will produce that particular effect.

When this fact finally sinks in, the problem of explaining purposive behavior returns with full force. We can no longer just wave vaguely at the details and say that one of the actions that could have had the observed effect must have occurred. We must account for the fact that the observed effect is produced and produced again, each time by THE ONLY ACTION THAT COULD HAVE PRODUCED IT AT THAT TIME.

If you think of the process of getting from one state of the environment to another, there are of course many trajectories by which this can happen even in a single instance of behavior. To understand the import of the point I'm trying to make you have to see the world at each instant, not over a series of instants. At any point during any action, the state of any variable including all its time derivatives is determined by the sum of all influences on it. If the trajectory is to repeat, then the action of the organism must at all times be exactly what is required to make up the difference between the trajectory that would have occurred without the influence of the action and the trajectory that is to be reproduced. And at every instant, there is only ONE action that can do this.

The basic problem in explaining behavior, therefore, is to explain how it is that the one action that is necessary to produce a given result is the one that is produced by the organism -- even when each instance of "the same behavior" requires that a specific different action be produced. This is the very question that Skinner's definition of "the operant" bypassed.

The other question raised by your comments concerns the role of the natural environment in behavior. Does the environment "make" organisms do anything? I claim it does not. But if it doesn't, what _is_ the part played by the environment in behavior?

One thing the environment does is to determine the physical effects of any action generated by the organism. The organism has, generally speaking, no way of altering the physical laws that determine these effects. If you exert a force on a free mass, it will accelerate; the only way you can keep it from accelerating is to remove the force or add a second canceling force. There is no way any organism can learn to apply a net force to a free mass without accelerating it.

A second thing the environment does is to have physiological and physical effects on the bodies of organisms. A human organism will die without oxygen, food, water, the right temperature range, and protection against physical damage, to mention a few items. There is nothing an organism can learn that will free it from such effects.

And a third thing the environment does is to stimulate sensory nerve endings. When we speak about the environment causing behavior, the causal path we usually mean involves the sensory systems. The question really is, can the environment make the organism produce certain actions by stimulating its sensory endings in the right way?

Let's not confuse this causal path with others. It's true that if the environment is suddenly depleted of oxygen, a person would "respond" by dying. So a requirement of continued life is a continued supply of oxygen. But this fact does not make the organism seek oxygen. If the organism did nothing to counteract the lack of oxygen, it would simply die. The question is not whether it is necessary for life that the environment be in a certain state; it is whether deviations from that state can, in themselves, make an organism do anything in particular.

The answer is, of course, no. The only possible way in which lack of oxygen could stimulate an organism to seek oxygen would be for the organism to possess sensory equipment that could report on the state of the oxygen supply, and internal organization that would convert a bad report into an action that would have the effect of restoring the oxygen supply. Whether such an action would occur does not depend on anything in the environment. It depends on processes inside the organism. If those processes are missing, the organism will die. If a lack of oxygen could make an organism behave to restore the oxygen supply, then organisms would never die from lack of oxygen. The environment would make them do what is necessary.

> A search for "bigger" reinforcers or for the CCEV could look very much alike.

Suppose we say that food is a reinforcer: more food, more behavior that produces food.

If we are looking at food as a CCEV, then we see it a little differently. We say, the more food there is, the less behavior there is to produce food. When the amount of food reaches a certain level, the behavior that produces it disappears altogether.

Are you sure you want to say that searching for reinforcers looks very much like searching for CCEV's?

It seems to me that this, not the points you cite, is the basic irreconcilable difference between EAB and PCT. We are talking about a FACTUAL difference. When the amount of reinforcement increases, does behavior increase (EAB) or decrease (PCT)?

In recent posts, I have suggested, and Bruce Abbott has tentatively agreed, that the apparent increase in behavior due to an increase in reinforcement that is seen under some schedules of reinforcement is actually due to the organism's turning to other kinds of behavior when the reinforcement rate is low, and spending more time on a particular behavior when the reinforcement rate associated with that behavior is higher than for other behaviors. The implication is that if the organism continued to be engaged in a particular reinforcement-producing behavior, the relationship would be that an increase in reinforcement goes with a decrease in behavior and vice versa -- the opposite of the basic assumption of EAB, but consistent with PCT.

> PCT then takes the position that events in the environment are only the ultimate cause of behavior overall because the subject would not even exist in the environment if some events had not occurred.

The cause of existence is different from the cause of behavior -- i.e., specific actions. Not so?

> If I have a reference for "picking up mount Shasta with my hands", the "environment" will "stop me" from doing such. To PCT the "consequences" of my attempt are neither a "reinforcer" nor a "punisher". PCT does recognize that somehow my "control failure experience" is "learned". PCT does NOT predict that I will immediately discontinue attempting to pickup mountains as a result of this failure or any other. THE MERE FACT OF THE EXISTENCE OF A CONTROL FAILURE IS NOT A GUARANTEE OF BEHAVIORAL CHANGE!

I think you're treating this on too intellectual a plane. If you picture actually trying to pick up Mount Shasta with your hands as a REAL PROCESS, I think you would discover a lot of reasons for reorganizing. Your hands would be raw and bleeding; you would have ruptured tendons and torn muscles all over your body; you would be in a state of chronic and complete exhaustion. I predict that you would very quickly give up this attempt because of extreme errors in other control systems, including intrinsic systems that are involved in reorganization. It's not the environment that stops you from having this goal; it's the felt consequences of trying to carry it out. If you didn't mind bleeding and aching and gasping for breath, nothing else could keep you from going right on with the attempt. The environment doesn't care how you feel, and it won't stop you from even suicidal efforts.

Best, Bill P.

Date: Mon, 26 Jun 1995 11:20:51 -0400
Subject: Operant psychology

[FROM: Dennis Delprato (950626)]

>Bill Powers (950621)

> There is far more involved in the Skinnerian interpretation than a mere listing of spatiotemporal processes over a period of time. There is a deep-seated _causal interpretation_ that has the same roots as do other ideas in psychology.

> The whole thrust of psychology as a new science was to interpret the behavior of organisms as the behavior of a mechanism like any other physical mechanism. I'm sure this orientation was consciously adopted very early in the history of psychology.... in the middle of the 20th Century when the automation revolution and the computer revolution began. Old concepts of <mechanism began to crumble as amazing discoveries were made about the <capacities of artificial computing devices and control mechanisms. The firm boundary between "mental" and "physical" phenomena began to waver, until now nobody who is knowledgeable about the new machines takes the distinction seriously.

Agree--you are describing the move away from lineal mechanistic to field/system constructions, as well as the associated rejection of substantive dualism. But see next.

> This is one reason, Dennis, that I see your firm insistence on "spatiotemporal" explanations as somewhat old-fashioned. It has the flavor of the old conception of mechanism, in which mental phenomena are considered ghosts in the machine rather than realities of the physical universe. What we have discovered about the capabilities of artificial computers and control systems has taught us that organization and function have a reality just as certain as the reality of matter and energy. Not only that, we have found that organization and function, while residing in physical matter, can transcend physical matter in the sense that the same organizations and functions can be realized in an endless variety of physical instantiations. The particular physical particles and physical laws that are used to implement computing or control functions are unimportant; every year more and different physical means are found for storing bits and performing logical functions. What matters is how the various physical elements are _related_ to each other, not what the physical elements are. And these relationships are not spatiotemporal in nature. They are functional and organizational -- what used to be called "mental."

The above confuses me (Dennis Delprato's views) with what I say in attempting to transmit Skinner's position (call me here, Del Dennis). What you have said just above gets at the heart of why I (Dennis D.) do not buy Skinner's approach. You have also transmitted something that gets at a fundamental conflict between Skinner's position and that of a sometime relative, J. R. Kantor. I have long taken Kantor's "interbehaviorism" as a significant departure from Skinner's radical behaviorism for reasons given above. Briefly put, with the field/system construct as a central key, inter-behaviorism is considerably more abstract than the straightforward Skinnerian accounts. [However, I will point out that very few today offer straightforward (orthodox) Skinnerian interpretations.]

As an example of what I am trying to say, not too long ago, a few of us knowledgeable about interbehaviorism and radical behaviorism had a discussion, it turns out, about relationships. I argued that relationships are abstractions, not observable. This was taken as something novel.

I insist that even though I know very little about the intricacies of control system modeling, I understand the "big picture" of PCT because of my work with inter-behaviorism. And, Skinnerians cannot comprehend the basic message of PCT because they BEGIN with a fundamentally incompatible postulate system, i.e., one tied to one-way cause-and-effect. This despite the fact that Skinner supposedly took a great step away from causes as instigators with "selection by consequences." Well, I am afraid this doesn't take them to 20th century science.

> A great part of Skinner's emphasis was on showing the superficiality of mental explanations of behavior, explanations that invoked unobservable entities and intervening variables. For every mental explanation, he tried to supply an equivalent explanation that involved only externally observable relationships among physical variables.

I do not find _relationships_ observable if we distinguish between observation and inference (admittedly, this is not a dichotomy). I suggest the mistake is to allow one's entire knowledge system to be based on the mistaken notion that when dealing with relationships between variables one is dealing solely with observables. This mistake will lead perpetuators to reject reference levels out of hand and to treat feedback as an external environmental stimulus, for example.

> The most frustrating (to me) aspect of Skinner's early work was not this interpretation, which was really standard among scientific psychologists, but the fact that he then proceeded to set up an apparatus in which a reinforcer would never appear unless the behavior occurred first. The sequence of events always began with a behavior, followed by a reinforcement. This causal relationship was right there in the environment, visible even to an untrained observer.

> Once an experiment got under way, however, reinforcements and behaviors occurred in a continuous stream, so as with the lawn ornament one could choose either way of perceiving causes and effects. As the animal worked toward mastery of the task, one could see the gradual increase in reinforcement rate as causing a gradual increase in behavior rate, or the gradual increase in behavior rate as causing a gradual increase in reinforcement rate. Once the rates had leveled out, one could see either the reinforcements maintaining the behavior, or the behavior maintaining the reinforcements. In the absence of any larger network of understanding of behavior, either interpretation would fit the observations.

Well put.

> Skinner, unfortunately, understood that the causes of behavior always lie in the environment. He therefore had no option but to see the reinforcement as the cause and the behavior as the effect, both during acquisition of the behavior and during maintenance of the behavior.

If one goes through enough of Skinner's corpus, one finds he offers conflicting statements on this issue, but, regardless of the details, he doesn't get it.

> So, Dennis, the historical record of which you speak does not speak for itself. It must be interpreted, and how it is interpreted makes all the difference.

I am not sure I follow this, but agree that constructs or interpretations cannot be avoided.

Dennis Delprato

Date: Mon, 26 Jun 1995 09:02:54 -0700
Subject: 600 pound E. coli

[From Rick Marken (950626.0900)]

Bill Powers (950624.0620 MDT) --

> It occurs to me that we've been missing our best bet for a task to compare PCT and reinforcement theory: a simple compensatory tracking task.

This task sounds fine but I think it is unlikely to have any more impact on reinforcement theorists than my E. coli experiment.

I was cut off from the net this weekend so I was unable to post the following analysis in a timely manner. But, as you will see, I have concluded that the results of the original E. coli experiment are as clear a rejection of reinforcement theory as you are likely to find. The fact that reinforcement theorists don't see this as a rejection of their model shows that there is nothing -- nothing at all -- that will convince reinforcement theorists that their model of behavior is wrong.

Here's what I wrote Saturday but was unable to post until today.

[From Rick Marken (950624...)]

Bill Powers (950624.1005 MDT) --

> In a real science, when a prediction fails you don't excuse it by saying that something unobserved must have happened in just the right way to keep the theory true. You investigate why it failed and produce the data that explains the failure.

> The basic question is whether behavior really works as each side says it works. Resolving that question requires thinking up tests that could give preference to one theory over the other. It requires taking failures seriously and using only observed data to explain them.

All sentiments which I applaud.

If reinforcement theorists were, indeed, playing science rather than defending the faith they would have seen that the results of the E. coli experiment gave a clear preference to PCT because they demonstrate the failure of a prediction of reinforcement theory.

I have re-read the original E. coli paper (p. 79 - 85 in _Mind Readings_) and see that the results show quite clearly that a prediction of reinforcement theory (selection by consequences) fails. The mistake I made on the net was to invite reinforcement theorists to explain the results of that experiment. This was a mistake because I understand reinforcement theory better than the reinforcement theorists. What the reinforcement theorists did is just what Bill describes above; they excused their failure to predict the E. coli results by saying that something unobserved must have happened in just the right way to keep their theory true.

What I did in the E. coli paper (which I didn't do on the net) was actually test reinforcement theory. I did this by measuring the reinforcement value of each of the consequences of responding. The reinforcement value of a consequences was measured as every textbook on reinforcement theory says it should be measured; as the tendency of a consequence to increase or decrease the probability of the response that produced it.

I measured the probability of a response as the inverse of the time to repeat the response that produced the consequence; the smaller the time until response repetition the higher the probability of the response. The reinforcement value of a consequence is then measured as the probability of the response that produces that consequence. The results looked like this for most subjects:

```
       |                                        x
       |                                   x
       |                             x
  P(R) |                      x
       |                x
       |         x
       |   x
        ---------------------------------------------
          0                                      180
          Consequence of R (Angle relative to target)
```

Figure 1

If the consequence of a response (bar press) is movement toward the target (0 degree angle relative to the direction of the target) then there is a very low probability of repeating the response; if the consequence of a response is movement away from the target (180 degree angle relative to the direction of the target) then there is a very high probability of repeating the response. So movement toward the target is the least reinforcing consequence of a response, movement away from the target is the most reinforcing consequence of a response, with reinforcement value varying linearly between these extremes.

Reinforcement theory says that the different directions of movement that follow a press will differentially strengthen the responses that produce them; since all directions of movement are equally probably, responding will be randomly reinforced.

So reinforcement theory predicts that the responding in this experiment will be random (this prediction was confirmed by running a reinforcement model --- including discriminative stimuli -- of the task; the reinforcement model responds randomly; the result is that the dot does a random walk around -- and off-- the screen).

The actual results of the E. coli experiment are not those predicted by reinforcement theory: responding is not random. Responding is actually quite systematic: the result of responding is NOT a random walk; it is control. The dot moves toward and remains near one of the three targets on the screen.

Reinforcement theorists (the reviewers of the E. coli paper) simply would not accept the notion that the results of this experiment are not what are predicted by reinforcement theory; they claimed that the results are predicted by reinforcement theory, but they never explained why.

Finally, Bruce Abbott came up with an ad hoc explanation of the E. coli results, ostensibly based on reinforcement theory. The explanation was based on the assumption that the reinforcing consequence is not the direction of dot movement after a response; it is _change_ (from before to after the response) in direction of dot movement.

What I didn't notice when Bruce proposed this explanation was that the choice of change in direction as the reinforcer was completely "extra-theoretical"; change in direction was selected as the reinforcer because it was just what was needed to keep reinforcement theory true.

In fact, change in direction of dot movement following a press would be REJECTED as a reinforcer by reinforcement theory because there is no differential reinforcement value of this consequence. The best way to see this is by noting that the probability of a response following a particular direction relative to the target (Figure 1) is the same regardless of the direction of movement prior to a response. The reinforcement value of moving, say, toward the target after a press is the same regardless of the direction of movement before the press; there is no differential reinforcement (measured as specified by reinforcement theory) based on change in direction of movement; only the direction of movement after the response affects the probability of response.

So reinforcement theory would have to predict that using _change_ in direction of movement as the reinforcer would produce the same result as using direction of movement: the prediction is a random walk IF _change_ in direction, like direction, is a random consequence of responding.

But using change in direction as the reinforcer does not produce a random walk! Why? The answer, of course, is "because change in direction after a response is NOT random". In fact, a change away from the target is more likely when moving toward it and a change toward the target is more likely when moving away from it.

So, by changing the definition of the reinforcer to a change in direction, reinforcement theorists played a trick and made it seem like their theory predicted a result that it does not predict. The trick was to call a consequence of responses a reinforcer when it was NOT a reinforcer; since this consequence is also non-random, the result of using this consequence as a reinforcer is non-random responding.

The E. coli demo was aimed at showing that people can respond systematically even though the reinforcing consequences of their responses are random. Reinforcement, which says that responding is selected by consequences, predicts that people will respond randomly if the reinforcing consequences of their responses are random. The results of this experiment clearly reject the predictions of reinforcement theory.

But the reinforcement theorists would not accept this failure of prediction; their theory MUST be true. So they used a trick to make it seem like reinforcement theory could produce systematic results when consequences are random. They did this by identifying a consequence of responding that is NOT randomly related to responses (change in direction after a response) and declaring that this consequence is the "real reinforcer" when, in fact, their own theory would REJECT it as a reinforcer.

Unfortunately, I fell for this trick and accepted the reinforcement theorists' pseudo-reinforcer (change in direction after a response) as a reinforcer. So I then showed that if change in direction after a response is made to be randomly related to responses (by biasing the probability of the direction of movement after a response) the reinforcement model again predicts random dot movement. Again, the actual result, with a human subject, is systematic responding, not the random responding predicted by reinforcement theory.

So even using the pseudo-reinforcer as the reinforcing consequence of responses, the prediction of reinforcement theory (random responding with random reinforcement) fails. To my knowledge, this failure was simply ignored by the reinforcement theorists.

The fact of the matter, however, is that the original E. coli experiment rejected reinforcement theory clearly and decisively.

The results of the E. coli experiment cannot be explained by reinforcement theory (without going outside the theory). However these results can be explained by a control model that acts in order to make a sensed representation of the results of its responses match its own specification for what those sensed results should be. This model can also explain all the results that reinforcement theory is supposed to be able to handle (such as the basic operant control task as well as control in the face of feedback function -- schedule -- changes).

The bottom line is that reinforcement theory says that responses are selected by their consequences; some consequences strengthen responses; other consequences weaken them. If this is true, then random presentation of strengthening and weakening consequences should produce random responding. Control theory says that organisms select consequences, not vice versa; responses are made in order to keep consequences in the reference state selected by the organism. Random changes in consequences are just a disturbance that will be resisted by responses.

When the reinforcement and control models of behavior were pitted against each other in the E. coli experiment, the reinforcement model failed; real organisms control consequences, they are not controlled by them.

The appropriate response to the E. coli experiment would be the abandonment of reinforcement theory in favor of testing the appropriate model of control -- perceptual control theory. Instead, behaviorists continue to defend a theory that has been decisively rejected. This is what PCT is up against. I don't know if there is any way to deal with people who simply will NOT accept demonstrations that their theory has failed.

Best Rick

Date: Mon, 26 Jun 1995 10:25:21 -0700
Subject: Well run behavior mod programs?!?

[From Rick Marken (950626.1025)]

Bruce Abbott (950625.1210 EST)--

> Well run behavior mod programs don't offer excuses for the failure of the manipulations to have their intended effects, they use systematic data collection to identify the source of the problem and find a way to correct it.

"Well run behavior mod programs"? This is an oxymoron, isn't it?

Happy 50th, by the way. I think you're old enough, now, to stop apologizing for reinforcement theory. The theory is scientifically wrong and humanistically repulsive. Yet many people still believe it and act according to its tenets -- to the detriment of civilized society.

How about spending the next 50 years making the world a better place; where everyone takes it for granted that people control, and are not controlled by, their environment.

Best Rick

Date: Mon, 26 Jun 1995 17:32:07 EDT
Subject: Re: 1200 dollar E. coli

[Martin Taylor 950626 17:15]
>From Rick Marken (950624...)

Rick, I'm afraid I must be quite obtuse. Obviously I don't understand reinforcement theory as well as Bruce, and you say:

> The mistake I made on the net was to invite reinforcement theorists to explain the results of that experiment. This was a mistake because I understand reinforcement theory better than the reinforcement theorists.

So take this as a naive comment/question.

I would have thought that making the cursor move faster toward, or less rapidly away from, the target would be reinforcing, and making it move faster away, or less rapidly toward, would be negative reinforcement.

If that is the case, then seeing the cursor moving away from the target, and pushing the button, would lead to reinforcement for button pushing more than it would lead to negative reinforcement, and if the cursor was seen moving toward, pushing the button would lead to negative reinforcement more often than positive. Both these ratios would depend on the angle:

```
          100 |                         x
              |                     x
   percent    |                 x
   positive 50|             x
 reinforcement|         x
              |     x
            0 | x
               ----------------------------
              0             90          180
                 degrees away from toward
                      target direction
```

The subject is heavily punished for pushing the button when the cursor is moving toward the target, heavily rewarded when it is moving directly away from the target, and punished and rewarded equally when it is moving at right angles to the direction to the target. Wouldn't this lead to the observed effects? The consequence of a button push when the cursor is seen moving to the target is usually bad, and of a button push when the cursor is seen moving away is usually good. So, selection by consequences would seem to yield more pushing when the cursor is moving away, and less when it is moving toward, the opposite of what you say it predicts.

I'm afraid I don't understand your analysis, and this naive approach may not be what "reinforcement theorists" would say. But it is what I would think of first, if someone asked me to describe a reinforcement-based description of the e-coli experiment.

I haven't observed too carefully the regions devoid of angel footprints, but I suspect this might be one of them.

Martin

Date: Mon, 26 Jun 1995 15:57:21 -0700
Subject: Re: 1200 dollar E. coli

[From Rick Marken (950626.1600)]

Martin Taylor (950626 17:15) --

```
>           100 |                         x
>               |                     x
>   percent     |                 x
>   positive  50|             x
> reinforcement |         x
>               |     x
>             0 | x
>                ----------------------------
>              0            90          180
>                 degrees away from toward
>                      target direction
```

> The subject is heavily punished for pushing the button when the cursor is moving toward the target, heavily rewarded when it is moving directly away from the target, and punished and rewarded equally when it is moving at right angles to the direction to the target.

Yes. That is what the empirical data show; in my graph the y axis was the empirically derived measure of the probability of a response following the consequence on the x axis.

> Wouldn't this lead to the observed effects?

It seems like it but no.

> The consequence of a button push when the cursor is seen moving to the target is usually bad, and of a button push when the cursor is seen moving away is usually good.

This is what fooled the reviewers. In fact, every consequence (from 0 to 180) is equally likely regardless of the current direction of movement relative to the target. So over all trials, the expected "percent positive reinforcement" following a press (regardless of the direction of movement before the press) is 50.

> So, selection by consequences would seem to yield more pushing when the cursor is moving away, and less when it is moving toward, the opposite of what you say it predicts.

That is ONLY true if the consequence of a press is defined as a CHANGE in direction. In that case, what you say is true. But CHANGE in direction after a press is not random in the sense that some changes (improvements) are more likely when you are going away from the target and other changes (deteriorations) are more likely when you are going toward the target.

> I'm afraid I don't understand your analysis, and this naive approach may not be what "reinforcement theorists" would say.

No. It IS what reinforcement theorists would say. That's how you can tell that they don't know their... never mind;-)

> But it is what I would think of first, if someone asked me to describe a reinforcement-based description of the e-coli experiment.

If you actually write the program that implements the reinforcement model you would see that the verbal analysis does not match the way the model actually behaves. That's why the reviewers never got it.

The reinforcement model works if you define reinforcement as the CHANGE in direction of cursor movement -- and assign positive reinforcing values to changes toward the target and negative or zero values to changes away from the target. This model works because the consequences of responding (changes in direction) are no longer (uniform) random. It is possible to make these change consequences uniform random by biasing the direction of movement following a press, the bias being based on the direction of movement before the press. When you do this, the reinforcement model behaves randomly (of course); random consequences select random responses. People in the same situation (consequences of presses are now random changes) do not produce random responses; they respond, as necessary, to keep the cursor on target.

Best Rick

Date: Mon, 26 Jun 1995 22:03:07 -0500
Subject: Ratio Model

[From Bruce Abbott (950626.2200 EST)]

Between starting summer school teaching and getting out book manuscript I've been a little bogged down today, but here's my two cents...

>Bill Powers (950623.1410 MDT) --

> OK, so you have to make the behavior rate depend on the value of the reinforcer to the subject, minus the cost of responding. How do you introduce the value of the reinforcer to the subject in the model? Does this require adding something to the model equivalent to a reference signal?

Yes, but a better equivalent would be measuring the relative size of the error signal. The "value" of a reinforcer basically is a measure of its motivating properties under given conditions. You can measure this value by, for example, determining the rate of responding that a given reinforcer will sustain under given conditions, e.g., a given quantity of a specified food under 12 hrs of food deprivation. A larger quantity, better taste, or greater hours of deprivation would be expected to yield higher rates, indicating that the reinforcer is more highly valued under those conditions. These operations increase the size of the error signal either by changing the reference level or by applying a larger disturbance.

>> Let's apply this model to the steady-state situation involved in performance on a simple CRF (1 response per reinforcement) schedule as ordinarily studied in the operant chamber. A hungry rat is placed in the chamber. We might conceive of the set point for rate of food pellet consumption as, say, 30 pellets per minute (30 ppm) under these conditions (essentially continuous eating). However, the apparatus limits the maximum rate to 10 ppm because of the delays involved in depressing the lever, moving to the food cup, picking up the food, devouring it, and returning to the lever. Thus, there is NO WAY that the rat can reach its reference level for this quantity, although it can reduce the error by a significant amount through lever-pressing (and by no other means). I believe that control theory predicts that the rat (once it has learned what to do) will develop a rate of responding that minimizes the error, up to the point where the error is reduced enough to bring the output below maximum. Depending on the gain, there will be a region within which the rate of responding will be a direct function of the magnitude of the error.

> All right, you're setting up a special situation in which the rat can't reach its reference level for nutritive input because there are physical limitations, combined with the reward size, that prevent its doing so. The exact prediction that a control model would make would depend on the loop gain of the rat (measured without these restrictions), on the measured reference level for food intake, and on the degree of the restrictions. If the rat had a high loop gain, we would expect the behavior to remain at the maximum rate under all conditions.

> Let's graph the situation:

```
>       450*
>          *
>          *     x
>          *
>          *          x
>       300* - - - - -  +
>BEHAVIOR  *            |  x
>          *            |
>          *            |       x
> RATE     *            |                      ref level
>       150*            |            x <--error--> |
>          *            |                          |
>          *            |                  x       |
>          *            |                          |
>          *            |                        x |
>          ************|*************|************|**********
>                       10            20           30
>                    REINFORCEMENT  RATE
```

> The line of x's represents the minimum consistent output sensitivity of the control system in units of presses per minute per reinforcement per minute of error. To fit your conditions it has to be about 30 (reasonable).

>> If we now increase the ratio requirement, the maximum rate of reinforcement available on the schedule is less (say, 5 ppm).

> It would have been easier if you had specified the new ratio requirement and then figured out the result, but we can still get there. If the new rate of reinforcement is now 5 reinforcements per minute, we can deduce that the new ratio is 390/5 = 78.

```
>       450*
>          *
>       390* - - x
>          *     |
>          *     |    x
>       300*     |
>BEHAVIOR  *     |         x
>          *     |
>          *     |              x
> RATE     *     |                             ref level
>       150*     |                   x <--error--> |
>          *     |                                 |
>          *     |                         x       |
>          *     |                                 |
>          *     |                               x |
>          ******|********************|************|**********
>                5     10             20           30
>                    REINFORCEMENT  RATE
```

I've been puzzling over these two graphs for several days now, trying to figure out where I'm getting lost. I think I now see the problem. Your graph has little to do with my scenario, and it's got an error in it to boot.

First, the error. If you are using an output sensitivity of 30, then in the first graph, the error between 30 rft/min and 10 rft/min is 30-10 = 20 rft/min, which would be expected to generate 30 * 20 = 600 responses/min and not the 300 you indicate in the figure. Right?

Second, the scenario. In my off-the-top-of-my-head example, I suggested that the maximum response rate on the CRF schedule is 10 rft/min, including the time required to collect the food, consume it, and return to the response lever. For the sake of illustration, let's assume that pressing and releasing the lever consumes 0.2 sec per press. Now, 10 rft/min is equivalent to one "cycle" of press-consume-return per 6 sec. If 0.2 sec is required for the press, this leaves 6.0 - 0.2 = 5.8 sec for the remaining activity. And at FR-1, 10 rft/minute = 1 * 10 = 10 ppm (presses per minute).

At the unspecified higher ratio, I stated that the maximum rate would be 5 rft/min, or 12 sec per cycle. With the consummatory activity still requiring 5.8 sec to complete, this leaves 12.0 - 5.8 = 6.2 sec for lever-pressing, or, at 0.2 sec per press, 31 presses. Thus (under the assumption of 0.2 sec/press) the new ratio is VR-31. At VR-31, 5 rft/min translates into 31 * 5 = 155 ppm (presses per minute). Remember, this is the _maximum_ rate at which food can be collected and eaten. So, in my hypothetical set of observations, the rat cannot reach its reference rate of food delivery on the CRF schedule (error = 20 rft/min) and does even worse under the new schedule (error = 25 rft/min), when responding at maximum rate. Yet there is an inverse relationship between rate of reinforcement and rate of responding, simply because of the extra time required to complete the ratio as the ratio requirement is increased.

Let's graph the situation:

```
          *
          *
       150*          +<---------------- error ----------//--->|
          *
          *
          *
          *
       100*
BEHAVIOR  *
          *                 +<--------------- error ----//--->|
          *
 RATE     *
        50*
          *
          *
          *
          *                         +<--------- error ---//--->|
          ************|*************|************|*******//****|
                      5            10           15            30
                     REINFORCEMENT  RATE
```

The error is increasing with the ratio requirement, but behavior is already at max rate and so can't get any faster. The line defined by the + symbols represents the constraint imposed by the time required to earn and collect the reinforcer, and not any regulatory effect.

Now, assume that the reference is 10 rft/min rather than 30, and that the output sensitivity is 30 as you suggest. On CRF the rat can earn its reference food rate by responding at maximum ppm. Assuming no constraints, at 5 rft/min the error is 10 - 5 = 5 rft/min, which translates into 5 * 30 = 150 ppm, giving a schedule value of 150/5 = VR 30. The maximum rate on this schedule is 152.5 ppm. Up to this point, at least, the schedule constraint does not enter in and the negative slope of the function is due to regulation.

I have more but it'll have to wait ... time's up.

Regards, Bruce

Date: Tue, 27 Jun 1995 00:37:34 -0400
Subject: Re: PCT and EAB

<[Bill Leach 950626.22:27 U.S. Eastern Time Zone]
>[From Bill Powers (950626.0630 MDT)]

> Probably the greatest hole in reinforcement theory was patched by the most casual plastering job. Skinner noticed way back in the beginning that what we casually describe as "the same behavior as before" is most often a very different behavior from the one we saw before. The actions

Which points out to me that I really did not understand the EAB meaning of "operant".

> The subject matter of PCT lies in the very part of behavior that Skinner dismissed as basically unexplainable. When he said that behavior is "emitted," he ceased to talk about the physical actions of the organism and began talking about _physical consequences_ of those actions.

and that can be soooooo close! (yet so far away)

> and over. We can't just say, as Skinner did, "Oh, well, the organism just did one of the many things that could have had the same physical effect." It's not as simple as that. In any one physical situation,

> ... What's really going on is that when the organism emits a particular physical effect, its action is PRECISELY THE ONLY ACTION THAT COULD HAVE PRODUCED THAT EFFECT AT THAT TIME. Given the same state of the environment, any other action would have resulted in a DIFFERENT effect.

> ... If you think of the process of getting from one state of the environment to another, there are of course many trajectories by which this can happen even in a single instance of behavior. To understand the import of the point I'm trying to make you have to see the world at each instant, not over a series of instants. At any point during any ...

I agree with this but it is doubtlessly related to where some of the misunderstanding comes from.

An individual repeating the "same act" under the same environmental circumstances will likely not employ exactly the same trajectory twice even if the final result is exactly the same.

Thus, we are faced with the fact that in most behavioral situations, a slight change in the output will still provide acceptable results. The set of correct actions vs the set of incorrect actions are both infinite sets of the same order. However if employing any discrete measurement system then there are many orders of magnitude more members in the incorrect action set.

I would suspect that "emitted behavior" type thinking fails to recognize that the theoretically infinite number of possibly correct trajectories completely masks recognition of the fact that typically an almost infinitesimally tiny variation with respect to the range of output in any one of the available degrees of output freedom is a "complete miss".

> oxygen and environmental paths

No, I definitely was not trying to say that lack of oxygen was any sort of control event. Of course many other "schools" of psychology don't see it as a control event either but then that might be because they don't recognize control at all.

Interestingly enough, there is no "natural" sensing of oxygen supply and thus this is a recurring "closed volume" cause of death problem.

>> A search for "bigger" reinforcers or for the CCEV could look very much alike.

> Suppose we say that food is a reinforcer: more food, more behavior that produces food.

I was being too kind! One very serious problem that I am having personally with "reinforcers" is that there seem to be at least two "kinds". When Bruce was talking about "rats preferring control" and Dennis; "Johnny wants attention", these seem like one kind and an object in the environment (ie: food) another kind.

Searching for a "reinforcer" related to "preferring control" should definitely look a lot like a CCEV search. Looking for a "reinforcer" similar to food (that is an object in the environment) would most certainly NOT resemble a PCT type search.

As mentioned before, the "ultimate cause in the environment" was meant to be a bit humorous. If a few events that occurred in the environment over which I had no possible control, I would not exist.

> intellectual plane

Yes, I agree. Such attempts at examples are anything but easy. I could add a few goals for no physical pain or injury in the attempt but the whole idea is just too weak. I can't think of a physical example that does not become too convoluted.

A miscellaneous comment:

We do have many years experience with engineered open loop "control" in well defined environments (the early "pattern" machine tools coming immediately to mind). These worked as long as the disturbances were not too great (and at that quite a bit of manual intervention was used).

OTOH, the _actions_ of a negative feedback machine tool are NOT exactly "predictable". The actual amount of movement and force varies as necessary to counter-act disturbance.

-bill

Date: Tue, 27 Jun 1995 10:57:26 -0500
Subject: Ratio Model, Continued

[From Bruce Abbott (950627.1055 EST)]

Continuing from where I left off in last night's post:

>Bill Powers (950623.1410 MDT) --
>>Bruce Abbott (950623.1135 EST)

>> But this analysis ignores response cost. Assume there is a second control system for "response effort" and that the reference for this perception is set to a low value. We now have two control systems in conflict when the rate-of-eating system brings the rate of lever-pressing up and thereby brings response effort above its set point. The result is that neither system can reach set point and rate of responding stabilizes at a kind of equilibrium value between the two set points.

> Actually, if you model this (as I have done), you will find that the only effect is a lowering of the effective responses-per-unit-error curve's slope. The slope does not reverse; it doesn't even become horizontal even if you make the cost very high. I was doing this in an effort to get the line of x-s to develop a curvature and go down again as the ratio requirement increased, to fit the Staddon/Motherall data. It doesn't even help to make the cost vary as the square of the response rate; the curve still never bends downward.

Yes, I can see that. Subtracting one linear function from another gives you another linear function. But who says that response cost is a linear function of response rate? I can run at top speed for about a block, but I can walk for miles.

> The only way to get the curve to bend downward is to use a cost-benefit control system as you suggest, but set its reference level so it doesn't come into effect until the error has reached about 60% of the maximum possible error. Then by adjusting the gain of this loop you can get the whole curve, over ratio requirements from 1 to 160, to match the real data very closely.

It seems to me that this solution is ad hoc. Why would one expect the reference level for effort to be set so high? More likely what you need, as suggested above, is a nonlinear input function in which response cost (as a perception, e.g., sensations of fatigue) remains low at low response rates but increases at an accelerating rate with increases in response rate. This kind of function seems to match known physiological effects of exercise rate. As the ratio requirement was increased, cost would increase at an increasing rate, thus moving the "compromise" set point of the two opposing control systems effectively closer to that of the effort system and producing a curved function.

>> You don't see the negative feedback regulation of food rate in the control model because food rate never reaches its reference value. For this reason I feel that your control model of the right portion of the Motheral curve does not apply in the way it would if control could be achieved.

> As the above curves show, there is negative feedback regulation throughout.

In last night's post, I tried to show that we were talking about different conditions and thus, different models.

> However, your example used unrealistic numbers. In the S/M data, the responses-per-unit-error slope was about 20. Increasing the ratio requirement from 1 to 2 decreased the reinforcement rate by about 5% (from 200 to 190) and doubled the response rate (from 200 to 380), roughly. Since the maximum observed response rate was about 3000, there is no question of being near or above the maximum response rate limit.

The Staddon/Motheral data related lever presses per session to dippers per session. The hypothetical data I presented gave the rates as lever presses per minute and reinforcements per minute, so the scales are different by a factor of 60 (assuming Motheral's sessions lasted one hour, which would be typical for an experiment of this type). The maximum response rate shown on the graph (I assume you're referring to Figure 7.18) is about 6000, which would translate to 100 responses/minute; this maximum was reached at 80% weight and 150 dippers/session, which gives a schedule value of VR-40. At 40 responses/dipper, this gives 2.5 dippers/min or 24 sec/dipper. If the rat requires 4 sec to visit the dipper and return to the lever, this leaves 20 sec to complete the response requirement, which allows 0.5 sec per response. Depending on the response-force and travel requirement of the lever Motheral used, this could be close to the maximum rate possible, but more likely it is not (I suspect something like 0.20-0.25 sec could have been obtained at maximum rate). So even at the higher rates generated under 80% weight it would appear that the maximum rate observed was not the maximum possible. So for Motheral's data, I agree with your analysis, that the right limb of the curve does not represent an effect of apparatus constraints as in my scenario, and that the rats were regulating the dipper rate.

So, even at 80% weight, the rats were not making an all-out effort to maximize the food rate. So much for half-starved rats. As I've noted before, the 80% weight criterion used in lab rats brings the rats' weights down to levels typical of those observed in rats in the wild and not to an emaciated state.

> I would prefer that in giving examples, you would use real data. Also, when you assume the effect of some added feature of the system, it would be a good idea to set up the model and run it, to see if your intuitive prediction of what would happen actually happens.

I'll do that. However, I would like to point out that in this case, my "intuitive" prediction of the effect I was trying to illustrate is right on the money, and there are experimental conditions under which it would be observed. Also, this prediction has nothing to do with reinforcement theory or PCT; it simply recognizes the effect of an apparatus constraint under certain conditions. What is true is that the scenario I envisioned does not fit the Motheral situation.

> If a reinforcement model CAN be made to fit the right side, I suggest strongly that you go ahead and develop it, however complicated it gets. The very complication, in comparison with the simplicity of the control model, may carry a message of its own. And if it should prove that you can't find any reinforcement model that, under the observed conditions and observed range of possible behavior rates, will fit the data, that will carry an even more important message.

Will do. But at the moment I'm having more fun working out a PCT analysis. Epicycles are boring. (:->

Regards, Bruce

Date: Tue, 27 Jun 1995 10:47:15 -0600
Subject: Re: ratio model

[From Bill Powers (950627.1012 MDT)]

Bruce Abbott (950626.2200 EST) --

You're right about my mistake; the scaling on my y axis is too small by a factor of 2 -- or else my assumed slope should be 15, rather than 30.

I think I understand your model now. The time taken per response is 0.2 sec and the minimum interreinforcement time is 5.8 sec.

The total time to produce m responses is 0.2*m/b, where m is the number of responses per reinforcement, and b is the number of responses per second. The units come out to seconds per reinforcement. Add to that the minimum time to retrieve one reinforcement, 5.8 sec per reinforcement, and we have

(1) r = 1/(5.8 + 0.2*m/b),

reinforcements per second as a function of behaviors per second.

The schedule provides a second relationship between r and b, namely r = b/m. The solution is where the curve of (1) crosses the line r = b/m. Substituting r = b/m into (1), we get

b/m = b/(5.8*b + 0.2*m) or

m = 5.8*b + 0.2*m, or

m(1 - 0.2) = 5.8*b, or

b = 0.8*m/5.8 = 0.1379*m

The behavior rate increases as the ratio requirement increases. Since r = b/m and b = m*r, we also get

m*r = 0.1379*m, or

r = 0.1379:

The reinforcement rate remains constant and independent of the ratio m.

I think you had better check this result before we go on!

Best, Bill P.

Date: Tue, 27 Jun 1995 13:02:22 -0700
Subject: Re: Ratio model, Continued

[From Rick Marken (950627.1300)]

Bill Powers (950623.1410 MDT) --

> If a reinforcement model CAN be made to fit the right side, I suggest strongly that you go ahead and develop it, however complicated it gets. The very complication, in comparison with the simplicity of the control model, may carry a message of its own. And if it should prove that you can't find any reinforcement model that, under the observed conditions and observed range of possible behavior rates, will fit the data, that will carry an even more important message.

Bruce Abbott (950627.1055 EST) --

> Will do. But at the moment I'm having more fun working out a PCT analysis. Epicycles are boring. (:-

I seriously doubt that reinforcement theory, even with "epicycles", can account for the ratio data any better than it can account for the E. coli data; that is, it can't.

Reinforcement theory can only account for the E. coli data by ignoring it. The E. coli data show that systematic responding occurs despite random reinforcement of responses. Reinforcement theory accounts for this result by ignoring the systematic behavior that occurs when reinforcement is random. If epicycle theory were like reinforcement theory, it would account for the retrograde motions of Mars by ignoring them.

Reinforcement theory says that reinforcing consequences increase the strength of behavior. Yet the ratio reinforcement experiments show that responding can be made to increase dramatically with no increases (indeed, with a slight decrease) in the amount of reinforcement that is presumably "maintaining" the behavior.

Before determining whether PCT is an improvement over the reinforcement model of the ratio data, it seems like it would be important to see that there IS a reinforcement model of this data.

Best Rick

Date: Tue, 27 Jun 1995 16:47:22 -0500
Subject: Ratio Model, Again

[From Bruce Abbott (950627.1645 EST)]

>Bill Powers (950627.1012 MDT)]
>>Bruce Abbott (950626.2200 EST)

> The reinforcement rate remains constant and independent of the ratio m.

> I think you had better check this result before we go on!

Good idea. Here's where the problem begins:

> I think I understand your model now. The time taken per response is 0.2 sec and the minimum interreinforcement time is 5.8 sec.

> The total time to produce m responses is 0.2*m/b, where m is the number of responses per reinforcement, and b is the number of responses per second. The units come out to seconds per reinforcement. Add to that the minimum time to retrieve one reinforcement, 5.8 sec per reinforcement, and we have

> (1) r = 1/(5.8 + 0.2*m/b),

> reinforcements per second as a function of behaviors per second.

The problem I was posing was, what happens when responding at maximum rate is not enough to bring the rate of reinforcement to its reference level? For the purpose of investigating this question I assumed that maximum response rate is 0.2 seconds per response (i.e., 5 responses/sec). If the reinforcement rate on the CRF schedule is 10 rft/min, this implies that the interreinforcement interval is 6 sec. With 0.2 sec of that interval taken up by the lever-press, that leaves 5.8 sec for collecting the reinforcer. Adding the ratio requirement m (responses per reinforcement) into the equation, we get

(1) t = 5.8 + 0.2*m seconds per reinforcement.

The number of reinforcements per second will be the inverse of this:

(2) r = 1/(5.8 + 0.2*m) reinforcements per second.

If you wish to produce an equation that gives the reinforcement rate as a function of the response rate (b), you must substitute 1/b for 0.2 in the above formula, giving

(2) r = 1/(5.8 + m/b) reinforcements per second.

This equation will hold until 1/b = 0.2, below which 1/b is a constant. This is equivalent to a rate of 5 responses/sec, the assumed maximum rate.

If the reference rate of reinforcement is very close to the maximum rate, there will be a region in which, as the ratio requirement increases, the resulting increase in error will drive the response rate up. When response rate reaches its maximum, that rate will be independent of error size, but interreinforcement rate will continue to decline because of the increasing time required to complete the larger and larger ratio requirements.

I've been looking for data in the literature on the relationship of behavior to the size of the ratio in variable ratio schedules. (It is surprising how little published data there are in which the ratio parameter was the specific subject of the investigation!) Here's one I think is interesting, although the study did not include enough different ratios (only three).

The study used condensed milk as the reinforcer, delivered via a dipper as in Motheral's study. Several dilutions were tested (more about that later). During initial (baseline) testing, the concentration was 30% and the subjects (4 rats at 80% ad lib weight) were exposed to VR-10, -40, and -80 schedules. The data reported include the overall response rate, the "running rate" (rate during completion of the ratio, excluding the "post reinforcement pause," and the duration of the post reinforcement pause (the time from delivery of the reinforcer until the next lever press). Here are the data for individual rats (estimated from the graph):

	VR-10	VR-40	VR-80	
Rat 1	0.8	1.5	1.8	responses/sec
Rat 2	0.8	1.4	1.7	
Rat 3	0.7	1.5	1.5	
Rat 4	0.8	1.5	1.5	

Average	0.8	1.5	1.6	responses/sec
	2880	5400	5760	responses/hour
	288	135	72	reinf/hour

```
Rsp/hour
 6000 *
      *       +
      *             +
      *
      *                 *
 4500 *
      *                     *
      *
      *                         *
      *
 3000 *                            +
      *
      *                                *
      *
      *                                   *
 1500 *
      *                                       *
      *
      * VR: 80     40             10             *  ref
      *      |      |              |                |
    0 **********|*********|*********|*********|*********|
      0        100       200       300       400       500
                          Dippers/Hour
```

The actual points are shown as plus signs in the graph. Using only the VR-10 and VR-40 points, the fitted line gives an output sensitivity of 16.5 and an estimated reference level of about 460 reinforcements/hour. Although it is dangerous to make much of the deviation of the VR-80 point from this line (given that the line is based on only two points), its position would be consistent with that portion of the curve which is beginning to level off prior to the downturn expected if we had the left limb of the complete function.

Now we come to some data we didn't have for Motheral's experiments:

	Running Rate (rsp/sec)				Post-Reinf Pause (sec)			
	VR 10	VR-40	VR-80		VR-10	VR-40	VR-80	
Rat 1	4.3	2.8	3.4		09	11	19	
Rat 2	3.5	2.6	2.5		08	11	11	
Rat 3	3.0	2.6	2.3		08	08	14	
Rat 4	4.6	2.6	2.7		08	08	16	
Average	3.9	2.7	2.7	rsp/sec	08	10	15	seconds
	234	162	162	rsp/hr				

It would appear that increasing the ratio requirement from VR-10 to VR-40 reduced the response rate during ratio runs but that a further increase of the ratio to VR-80 has little or no additional impact. This would be consistent with the rats maintaining a somewhat reduced running rate during longer ratio runs in order control the sensory effects of overexertion (long ratios at the higher rate might lead to oxygen starvation of the muscle or the buildup of lactic acid, for example).

However, increasing the ratio from VR-10 to VR-40 had little or no impact on the length of the postreinforcement pause but a further increase to VR-80 was associated with a lengthening of the pause. This latter effect would be consistent with a increase in time "doing other things" as the cost of earning the dippers of milk begins to exact too high a penalty for the benefit returned. (Note: The article states that the time required to consume the reinforcer is no more than 3 seconds.)

This is, of course, speculation, and I may be over-interpreting the data, which may not be reliable enough to trust the reality of the relationships indicated by the averages given above (see exceptions for individual rats).

During a subsequent test phase, the concentration of the milk was varied within-session from 10% to 75% and the duration of the postreinforcement pause was recorded FOLLOWING delivery at each concentration. Higher concentrations produced longer postreinforcement pauses, and the increase in pause with increased concentration was larger the higher the ratio requirement (although the data at VR-80 are somewhat messy). If higher concentrations are more effective in temporarily reducing the error in nutrient intake, this is what would be expected under the PCT model.

Here's the reference:

Priddle-Higson, P. J., Lowe, C. F., & Harzem, P. (1976). Aftereffects of reinforcement on variable-ratio schedules. _Journal of the Experimental Analysis of Behavior_, _25_, 347-354.

Regards, Bruce

Date: Tue, 27 Jun 1995 17:38:31 -0600
Subject: Re: ratio model

[From Bill Powers (950627.1415 MDT)]

Bruce Abbott (950626.2200 EST) --

It's my day for mistakes. Let me start over and take smaller steps.

The time between reinforcements is the eating time delay d plus the time taken to execute m presses at b presses/min.

t = d (min/reinf) + m (press/reinf) / b (press/min)
  = (d + m/b) min/reinf

d = eating time delay per reinforcement, min
b = behavior rate when active, presses per minute
m = ratio setting of schedule

The role of the 0.2 sec per press that you mentioned is simply to set the maximum possible behavior rate at 5 per sec or 300 per minute.

So the reinforcement rate is

(1) r = 1/t = 1/(d + m/b)

This is the modified schedule of reinforcements expressed as reinforcements per minute per press per minute, where presses per minute are measured only during intervals of continuous pressing.

To get a system of equations we need the organism-equation which expresses the actual behavior rate as a function of reinforcement rate. The PCT and reinforcement models, I think, would be different, as the reinforcement model assumes (roughly) b = k*r, while the PCT model assumes b = k*(r' - r), r' being the reference level.

I'll work on this some more in the morning. Feel free.

Best, Bill P.

Date: Tue, 27 Jun 1995 21:04:39 -0700
Subject: Re: Ratio model, again

[From Rick Marken (950627.2100)]

Bruce Abbott (950627.1645 EST) describes data collected by:

> Priddle-Higson, P. J., Lowe, C. F., & Harzem, P. (1976). Aftereffects of reinforcement on variable-ratio schedules. _Journal of the Experimental Analysis of Behavior_, _25_, 347-354.

Very interesting data, Bruce. But could you explain "running rate" in a little more detail. It sounds like "running rate" is the response rate that occurs after the animal has worked through the ratio requirement. Is that it?

Also, what were these researchers trying to find out by collecting these data? I presume they were testing some predictions of a reinforcement model? Did they compare their data to the data produced by a working reinforcement model? Were these data consistent with the predictions their model?

Best Rick

Date: Wed, 28 Jun 1995 03:59:14 -0600
Subject: Re: ratio model

[From Bill Powers (950628.0400 MDT)]

Bruce Abbott (950627.1645 EST) --

More good data, there. Did the authors make any comment at all about the fact that more reinforcement goes with less behavior?

I finally got the formula right. At the end there is some source code that plots the relationship between reinforcement rate and running behavior rate as a function of m, the ratio. The assumed fixed delay is initialized to 8 seconds as in some of the data you sent me.

The reinforcement model says that a higher rate of reinforcement maintains a higher rate of behavior, so I've shown that model over the schedule curves in red for various proportionality factors k. The control model says that the rate of behavior is determined by some reference rate of reinforcement minus the actual rate of reinforcement; I've plotted that in green for one value of gain, -21.

Clearly, for any value of the assumed k for the reinforcement model, the reinforcement rate has to rise with the behavior rate (the intersections of any red line with the white curves). There's also a problem with existence of a solution for many values of m, other than zero behavior rate and zero reinforcement rate.

The control model shows behavior decreasing as reinforcement increases, and there's no problem with existence of solutions, obviously.

The rat data in the second set was interesting.

	Running Rate (rsp/sec)				Post-Reinf Pause (sec)			
	VR 10	VR-40	VR-80		VR-10	VR-40	VR-80	
Rat 1	4.3	2.8	3.4		09	11	19	
Rat 2	3.5	2.6	2.5		08	11	11	
Rat 3	3.0	2.6	2.3		08	08	14	
Rat 4	4.6	2.6	2.7		08	08	16	
Average	3.9	2.7	2.7	rsp/sec	08	10	15	seconds
	234	162	162	rsp/hr				

This is clearly rather extreme behavior: 14000 responses per hour! I don't know what your last line above means -- it's obviously not 3600 times the line above it.

I would guess that this set of runs represents a very dilute solution of condensed milk; the curves seem to be on the border between positive and negative slopes. My hypothesis is that you get a positive relation between reinforcement and behavior when not enough reinforcement is received to support life if continued indefinitely.

There is one problem with using dilution as a way of altering the amount of reinforcer. The less the food content, the more water the animal has to drink to get the same amount of food reinforcer. So the water-loading control system is going to have a "too-much" error, and start fighting the food control system.

I think that when we analyze data we should do it on a rat-by-rat basis, not average the data across rats. In the above data set, the rats don't all behave alike. Better to apply the model to each rat, and then compare the parameters.

It's interesting that using variable ratio schedules seems to give the same kind of results that Motherall got with fixed ratios.

If you multiply the reinforcements/hour by the ratio in the above data, you get the average responses per hour instead of the peak rates. Motherall's data was in terms of average response rates.

I should be in bed.

Best, Bill P. [program followed. deleted here]

Date: Wed, 28 Jun 1995 09:31:45 -0500
Subject: Re: Ratio Model

[From Bruce Abbott (950628.0900 EST)]

>Bill Powers (950628.0400 MDT) --
> Bruce Abbott (950627.1645 EST)

> More good data, there. Did the authors make any comment at all about the fact that more reinforcement goes with less behavior?

The inverse relationship between the overall response rate and the schedule parameter has also been observed on random-ratio schedules (Brandauer, 1958, Kelly, 1974). These functions would appear not to be consistent with the Law of Effect (Herrnstein, 1961; 1970), as the Law would predict an direct and not an inverse relationship between response rate and rate of reinforcement, as determined by the schedule parameter.

(Priddle-Higson, Lowe, & Harzem, 1976, p. 353)

However, these authors felt that these results might be explained by the complex interaction between running rate and postreinforcement pause observed in the data. The overall rate would result from the combination of these two factors.

```
> Average   3.9     2.7     2.7  rsp/sec  08    10    15 seconds
>           234     162     162  rsp/hr
```

> This is clearly rather extreme behavior: 14000 responses per hour! I don't know what your last line above means -- it's obviously not 3600 times the line above it.

Sorry, I had been working with rsp/min and later converted to rsp/hr for comparison with Motheral's data. I forgot to do the conversion on the running rate data, as I did not plot them. So the numbers in the lower line are rsp/min. 14000 rsp/hr is a high rate, but that's not the sustained output over the session. On ratio schedules, once the rat begins the ratio, it nearly always knocks it off in one continuous series at a steady (and fairly high) rate.

The running rate reflects the time required to complete the ratio once the rat has resumed responding following the postreinforcement pause and is not the rate maintained overall. In my hypothetical scenario I had suggested 0.2 sec per response as a reasonable figure for the maximum rate possible, which gives 5 responses per second, compared with the 3.9 shown in these data. The highest running rate shown in the data presented in the article is around 5.2 rsp/sec. This rate was recorded for one animal on the VR-10 schedule at a 10% milk concentration, the lowest tested.

> I would guess that this set of runs represents a very dilute solution of condensed milk; the curves seem to be on the border between positive and negative slopes. My hypothesis is that you get a positive relation between

reinforcement and behavior when not enough reinforcement is received to support life if continued indefinitely.

The condensed milk was diluted to a 30% concentration, or about 2 parts water and 1 part concentrate. This is not especially dilute. The dipper presented 0.05 ml per dip, a medium-sized raindrop.

That's my hypothesis, too. However, you do have to keep in mind the fact that these rats will be given supplemental food after each experimental run, so a large deficit is not going to appear over the long run as it would if all the rat's nutrient intake had to come from lever-pressing. Given the actual arrangement, it is possible to induce a rat to work at a (short-term) loss by increasing the ratio requirement slowly enough, although even then there are limits beyond which the behavior will break down.

> There is one problem with using dilution as a way of altering the amount of reinforcer. The less the food content, the more water the animal has to drink to get the same amount of food reinforcer. So the water-loading control system is going to have a "too-much" error, and start fighting the food control system.

Yes, that could be a problem; on the other hand the stomach loading is constant on a per-reinforcer basis when the amount is varied in this way, so there are advantages as well as disadvantages to this approach. The water-loading control system has a simple way to compensate for the error other than reducing the intake rate: excrete the water faster. So perhaps this is not as serious a problem as it may seem to be at first blush.

> I think that when we analyze data we should do it on a rat-by-rat basis, not average the data across rats. In the above data set, the rats don't all behave alike. Better to apply the model to each rat, and then compare the parameters.

I agree, and that is why I presented the individual data in my post. One of the difficulties in gathering data like these is that it is not always possible to keep variables such as motivational level (error) perfectly constant over the course of an experiment. For this reason some of the individual data may not be as consistent as one would like. In the data presented, most subjects showed similar changes in rates across the schedules, but there were exceptions. The averages I graphed seemed to be representative of what most subjects did most of the time.

> It's interesting that using variable ratio schedules seems to give the same kind of results that Motherall got with fixed ratios.

Yes, although I'm not completely certain that Motheral used fixed ratios (Staddon keeps referring to "ratio schedules" without designating whether they were fixed or variable ratios). But if so, the similarity of results with fixed and variable ratios would have emerged despite differences between the two schedules in the local patterns of responding associated with them.

I haven't taken a look at your program yet, so I won't comment on it now.

>Rick Marken (950627.2100) --

> Very interesting data, Bruce. But could you explain "running rate" in a little more detail. It sounds like "running rate" is the response rate that occurs after the animal has worked through the ratio requirement. Is that it?

It's the response rate observed WHILE the animal is working through the ratio requirement, timed from the first response following the postreinforcement pause.

> Also, what were these researchers trying to find out by collecting these data? I presume they were testing some predictions of a reinforcement model? Did they compare their data to the data produced by a working reinforcement model? Were these data consistent with the predictions their model?

In the experiment I described the researchers were assessing the effect of milk concentration on the length of the postreinforcement pause, on the running rate, and on the overall rate of responding, in the interaction of these effects with FR ratio size. The study does not appear to have been designed to test predictions of a reinforcement (or any other) model, although there was some discussion of the anomaly apparent in the data with respect to reinforcement theory predictions. Rather, it was designed to provide information about the relationships being examined (Baconian style). They did not develop a working model.

Regards, Bruce

Date: Wed, 28 Jun 1995 09:00:58 -0700
Subject: The Naked Emperor

[From Rick Marken (950628.0900)]

Bill Powers (950628.0400 MDT) --

> More good data, there. Did the authors make any comment at all about the fact that more reinforcement goes with less behavior?

Bruce Abbott (950628.0900 EST) quotes Priddle-Higson, Lowe, & Harzem

> These functions would appear not to be consistent with the Law of Effect (Herrnstein, 1961; 1970), as the Law would predict an direct and not an inverse relationship between response rate and rate of reinforcement, as determined by the schedule parameter.

and adds:

> However, these authors felt that these results might be explained by the complex interaction between running rate and postreinforcement pause observed in the data.

This may be why Bill Powers once suggested that there might be a conspiracy among EABers to keep quiet about the failure of reinforcement theory. Here we have data that seems to be completely inconsistent with reinforcement theory (an inverse relationship between response rate and rate of reinforcement). And how do reinforcement theorists deal with what this monumentally important observation? They "feel" that it can be "explained" by the complex interaction between running rate and postreinforcement pause observed. Can it be explained by these observations? Apparently, nobody cares because reinforcement theory MUST be right.

I really don't see how the "complex interaction between running rate and postreinforcement pause" could possibly explain the inverse relationship between response rate and rate of reinforcement anyway. After all, whatever complex interaction between running rate and postreinforcement pause exists, it exists right along with the finding of the inverse relationship between response rate and rate of reinforcement. Perhaps the authors meant that this interaction can "explain away" rather than explain the inverse relationship; perhaps the hope was that by removing the variance due to the interaction, the unpleasant inverse relationship between response rate and rate of reinforcement would magically disappear.

It's amazing that there has never been ONE little boy in the EAB community - - in the entire psychological community-- who was willing to say that the Emperor of reinforcement has no clothes.

Best Rick

Date: Wed, 28 Jun 1995 10:57:40 -0600
Subject: Re: ratio model

[From Bill Powers (950628.0945 MDT)]

Bruce Abbott (950628.0900 EST) --

Sounds as though there is some precedent for commenting on the reversal of the relationship expected under the Law of Effect. When the authors say

> "These functions would appear not to be consistent with the Law of Effect (Herrnstein, 1961; 1970), as the Law would predict an direct and not an inverse relationship between response rate and rate of reinforcement, as determined by the schedule parameter."
>
> However, these authors felt that these results might be explained by the complex interaction between running rate and postreinforcement pause observed in the data. The overall rate would result from the combination of these two factors.

This is a pretty feeble comment on an observation that goes directly against the fundamental assumptions behind reinforcement theory itself. When you compute the _average_ values of reinforcement rate and behavior rate, the same relationship is seen. In fact, the equivocal slopes become the "wrong" slopes for reinforcement theory. (You can get the average behavior rate by multiplying the average reinforcement rate by the ratio).

When you run my program, I think you will see that the "complex interaction" isn't all that complex and is certainly not sufficient to explain the result.

The attitude displayed by the authors toward this finding is interesting from the standpoint of a study in steadfast faith. "No, madame, there is no danger. A little bit of floating ice could never damage the Titanic." I wonder what they will say when they discover that this reversed relationship can be found in _every_ experiment in which the ratio schedules cover a wide enough range. It's even there in Skinner's descriptions of how he shaped pigeons to a high rate of responding. He did it by _reducing_ the amount of reinforcement per peck.

> In the experiment I described the researchers were assessing the effect of milk concentration on the length of the postreinforcement pause, on the running rate, and on the overall rate of responding, in the interaction of these effects with FR ratio size.

That is what they _thought_ they were assessing. Actually, they were simply looking at the _relationship_ between milk concentration and those other variables. To say they were assessing the _effect_ of milk concentration is to assume that milk concentration is _a priori_ a causal variable.

RE: water loading

> The water-loading control system has a simple way to compensate for the error other than reducing the intake rate: excrete the water faster.

I think we have to distinguish "appetitive" control systems, which control the immediate effects of ingested food or water, from "chronic" control systems, which work on a much slower time scale and control longer-term average effects. If you load the stomach with 0.05 ml of water every 8 seconds or so, that is almost half a milliliter per minute or about 20 ml per hour -- a lot for a rat's stomach. If this "simple way to compensate" is seriously offered as reason to ignore the confounding of food intake with water intake, then the experiment should be done to verify that this is justified.

> One of the difficulties in gathering data like these is that it is not always possible to keep variables such as motivational level (error) perfectly constant over the course of an experiment. For this reason some of the individual data may not be as consistent as one would like.

A more likely explanation in my view is that there are individual differences between rats. While an individual rat might behave in a way perfectly consistent with a model, the appropriate parameters might be quite different from one rat to another.

I don't think we would want to keep a rat's "motivation level (error) constant" in an experiment, even if we could. If you want to measure the characteristics of a control system, you have to let it control, and this means letting it correct its errors. The only way to keep error constant would be to use an external control loop that is stronger than the rat's control system; this would not reveal normal control characteristics and might well cause reorganization to start, making all measures invalid.

> In the data presented, most subjects showed similar changes in rates across the schedules, but there were exceptions.

The greatest percent differences were in the data set where the peak rates of responding were near 14000 per hour, and the average rates around 6000. That is in the upper left corner of the data plots where the slope of the relationship may be on the point of reversing. This is not the best place to evaluate the normal relationships among variables.

> I'm not completely certain that Motheral used fixed ratios (Staddon keeps referring to "ratio schedules" without designating whether they were fixed or variable ratios).

Now that I look at the book again, I'm not sure, either. I can't find any specification as to the type of ratio schedule. My model, which uses fixed ratios for the schedule, fits the data. But I suspect that a variable ratio schedule would work, too. There's no basic difference except for the noise level in the data.

> The study does not appear to have been designed to test predictions of a reinforcement (or any other) model, although there was some discussion of the anomaly apparent in the data with respect to reinforcement theory predictions.

Whether it was designed to do this or not, it was a test of the reinforcement model. Every experiment is a test of the basic model, isn't it?

Best, Bill P.

Date: Wed, 28 Jun 1995 20:24:31 -0500
Subject: Re: Ratio Model

[From Bruce Abbott (950628.2020 EST)]

>Bill Powers (950628.0945 MDT) --
>>Bruce Abbott (950628.0900 EST)

>> However, these authors felt that these results might be explained by the complex interaction between running rate and postreinforcement pause observed in the data. The overall rate would result from the combination of these two factors.

> This is a pretty feeble comment on an observation that goes directly against the fundamental assumptions behind reinforcement theory itself.

Well, these guys aren't reinforcement theorists, they're just humble experimentalists reporting their data. I give them credit for perceiving and pointing out the apparent problem for reinforcement theory; apparently they felt less than comfortable going much further. Their objective was to discover what relationships occur in this situation, and that's what they did. They were willing to leave the theorizing to someone else.

> When you compute the _average_ values of reinforcement rate and behavior rate, the same relationship is seen. In fact, the equivocal slopes become

the "wrong" slopes for reinforcement theory. (You can get the average behavior rate by multiplying the average reinforcement rate by the ratio).

Huh? The _same_ relationship when you compute the average values? The average, overall rates are precisely the ones the researchers were talking about when noting the inverse relationship, so what other relationship can you be speaking of? I'm confused.

What the researchers were noting is that the overall response rate can be viewed as being determined by two factors: the running rate and the postreinforcement pause, and that these changed in different ways with the ratio requirement. In some cases these changes tended to oppose one another in their effect on the overall rate; in others they tended to summate. Given that the overall changes reflect an average of these more local changes, then one might suspect that the overall rates obfuscate the actual processes at work. That seems reason enough to be at least a little careful about announcing to the world that your data absolutely contradict a reinforcement analysis. If you found some rather odd things happening in a complex tracking task that seemed to contradict PCT, you might note that the results don't appear to agree with the PCT prediction but I doubt you'd be ready to chuck the whole theory.

> The attitude displayed by the authors toward this finding is interesting from the standpoint of a study in steadfast faith. "No, madame, there is no danger. A little bit of floating ice could never damage the Titanic."

No, it's more like "there's something odd going on here, but the details are rather complex so I'm not exactly sure how to interpret these results. I'll leave that to someone else."

>> In the experiment I described the researchers were assessing the effect of milk concentration on the length of the postreinforcement pause, on the running rate, and on the overall rate of responding, in the interaction of these effects with FR ratio size.

> That is what they _thought_ they were assessing. Actually, they were simply looking at the _relationship_ between milk concentration and those other variables. To say they were assessing the _effect_ of milk concentration is to assume that milk concentration is _a priori_ a causal variable.

Oh, come ON! If I can't say that a disturbance to a control system under given conditions _causes_ it to respond with an opposing action, then I don't know of _any_ circumstances in which the term would be appropriate. There is directionality to this relationship; certainly the changes in response rate, reinforcement rate, and postreinforcement pause did not cause the milk concentration to change. The word "cause" implies this directionality in a way that "relationship" does not. You may be able to rearrange F = MA or I = E/R and get sensible results, but this is different.

>RE: water loading

>> The water-loading control system has a simple way to compensate for the error other than reducing the intake rate: excrete the water faster.

> I think we have to distinguish "appetitive" control systems, which control the immediate effects of ingested food or water, from "chronic" control systems, which work on a much slower time scale and control longer-term average effects. If you load the stomach with 0.05 ml of water every 8 seconds or so, that is almost half a milliliter per minute or about 20 ml per hour -- a lot for a rat's stomach. If this "simple way to compensate" is seriously offered as reason to ignore the confounding of food intake with water intake, then the experiment should be done to verify that this is justified.

I agree. But I'm sure these investigators kept a record of the response rates throughout the session and would have reduced the size of the droplet or the session length if there were evidence of satiation. Perhaps that's why they started with VR-10 rather than something lower. As to time-scale, I would bet that a rat's kidneys can do quite a bit of error reduction within the span of an hour.

>> One of the difficulties in gathering data like these is that it is not always possible to keep variables such as motivational level (error) perfectly constant over the course of an experiment. For this reason some of the individual data may not be as consistent as one would like.

> A more likely explanation in my view is that there are individual differences between rats. While an individual rat might behave in a way perfectly consistent with a model, the appropriate parameters might be quite different from one rat to another.

If parameter adjustments can allow a single model to account for the conflicting changes seen across ratios in some of these animals, then I'm worried about the model--it would appear to be able to account for any pattern whatever. When you have that, you're doing curve fitting.

> I don't think we would want to keep a rat's "motivation level (error) constant" in an experiment, even if we could. If you want to measure the characteristics of a control system, you have to let it control, and this means letting it correct its errors. The only way to keep error constant would be to use an external control loop that is stronger than the rat's control system; this would not reveal normal control characteristics and might well cause reorganization to start, making all measures invalid.

Depends on which level you're talking about. Generating a given level of error in the nutrient-level control system simply establishes a given reference level for lower-level systems such as the one governing rate of eating. For the purpose of studying the relationship between lever-press rate and eating rate as the ratio requirement is varied this may be perfectly acceptable procedure. This is the system I'm measuring the characteristics of, and I'm letting it control.

> The greatest percent differences were in the data set where the peak rates of responding were near 14000 per hour, and the average rates around 6000. That is in the upper left corner of the data plots where the slope of the relationship may be on the point of reversing. This is not the best place to evaluate the normal relationships among variables.

Not in the data I'm looking at. The peak (running) rates near 14000 per hour occurred on the VR-10 schedule, which is at the lower right end of the curve. The peak (average) rate (5760) occurred on the VR-80 schedule, which is the leftmost point, but this rate is only slightly higher than the rate observed on VI-40 (5400).

>> The study does not appear to have been designed to test predictions of a reinforcement (or any other) model, although there was some discussion of the anomaly apparent in the data with respect to reinforcement theory predictions.

> Whether it was designed to do this or not, it was a test of the reinforcement model. Every experiment is a test of the basic model, isn't it?

Yes, but not every experiment yields an unambiguous test of theory, and such ambiguity is especially likely when the experiment was not designed with theory testing in mind. Some studies are just carried out to discover empirical relationships (as I think this one was). After the data have been collected and published, the theorists can then step in and squabble over the implications. (;->

Regards, Bruce

Get CSG1 & 2 as of July 1 am

[Though this discussion will continue, this thread will end here. Future editions of PCTtexts will require more file space -- more disks.]