Causal Dimensions that Create Difficulty in Understanding Evolution

Michael Edgar and Tina Grotzer
Harvard University
Presented at the National Association of Research
in Science Teaching (NARST)
New Orleans, April 28- May 1, 2000

The Understandings of Consequence Project
Project Zero, Harvard Graduate School of Education
125 Mount Auburn Street, 5th Floor
Cambridge, MA 01238

This paper is based on the results of research carried out during the first year of the Understandings of Consequence Project. We are continuing to research and develop the ideas presented here. If you have feedback for us or would like to keep in touch with developments on the project, please check our website at http://pzweb.harvard.edu/Research/UnderCon.htm or send us an email at Tina_Grotzer@PZ.Harvard.Edu or Michael_Edgar@PZ.Harvard.Edu

This paper is based upon the work of Understandings of Consequence Project, which is supported by the National Science Foundation, Grant No. REC-9725502 to Tina Grotzer and David Perkins, Co-Principal Investigators. Any opinions, findings, conclusions or recommendations expressed here are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Introduction

“Why has Darwin been so difficult to grasp?” ponders Steven Jay Gould in the prologue to his 1977 book, “Ever Since Darwin.” Gould suggests that the difficulty lies not in the scientific difficulty but in the radical philosophical switch required by the theory. While the basic tenets seem straightforward and simple enough, Darwin asks us to make the following, more difficult assumptions according to Gould: 1) Evolution is purposeless. Any resulting harmony is a consequence of individuals seeking their own genetic advantage, not of some grand scheme; 2) Evolution is directionless. It is not leading to some better, higher state. Rather it leads to adaptedness to local environments; and 3) Matter is the stuff of all existence. Wondrous things such as minds, spirit, and even God are the result of neuronal complexity. This physical materialism, as Gould calls it, is difficult for us to grasp and also is at the root of many misgivings about evolution. Gould has helped us to see how the implications of Darwin’s “simple” tenets can deeply challenge our basic beliefs about the world. In this paper, we expand upon this line of reasoning to explore further cognitive and basic philosophical challenges that understanding evolution and the process of natural selection poses.

We argue here that understanding evolution is cognitively challenging because it contradicts many of our default assumptions about the nature of cause and effect. Certainly, as Gould suggests, the implications of natural selection also challenged philosophical beliefs of the time. However, even persons who claim to believe in evolution and are some its more ardent supporters have difficulty grasping and holding onto its tenets (e.g. Lord & Marino, 1993; Zimmerman, 1987). Zimmerman (1987) found in a survey of one thousand high school biology instructors, that approximately 25% of them held misconceptions about the theory, for instance that it involved the purposeful striving of an organism towards a higher more complex form of life–a Lamarckian conception.

Given such difficulties grasping evolution broadly and natural selection specifically, it behooves us to ask whether there are other aspects of the theory that pose specific cognitive challenges for people. We hypothesize here that there are at least ten persistent difficulties that people have and that these may be due in part to assumptions that people make in their causal reasoning and in particular, how causes and effects behave. For instance, people often seek out centralized control structures as the causal mechanisms behind “how things are” and yet, the theory of evolution asks us to envision a decentralized causal mechanism, where outcomes emerge from a multitude of interacting events that share “control” for the outcome in a distributed fashion (Resnick, 1996). The theory of evolution is used to explain the magnificence of the life forms around us and yet the causal mechanism of natural selection is unintentional with no aim in the generation of life forms and only an aim in the selection of life forms–those that fare best. These ways of thinking can be counterintuitive even to those who accept

Darwin's tenets.

Previous Research on Core Causal Intuitions

The idea that people bring specific causal expectations to their thinking and learning has been argued by others and has been demonstrated in other areas of scientific understanding. For instance, di Sessa (1993) considered whether there are notions of causality underlying alternative conceptions in physics. He introduced the concept of phenomenological primitives (or p-prims), small knowledge structures that people use to describe a system's behavior. These schemata come into play as ready explanations or components of explanations. They are often considered to be self-explanatory and to need no justification. Di Sessa does not view them as part of a unifying theory. He considers them to "belong neither to the lowest, possibly 'hard-wired' and data driven sensory elements, nor to the world of ideas, named concepts or categories" (1993, p. 112). These p-prims behave as defaults for how people expect systems to behave. They are rarely considered in a reflective sense. Similarly, Brown (1996) refers to core causal intuitions that lead students astray as they bring them to bear in attempting to understand electricity concepts. He focuses on core intuitions regarding how people attribute agency and how they assess responses to agency. He identifies a number of types (initiating; initiated; reactive and so on.) Andersson (1983) draws upon Lakoff and Johnson's (1980) notion of an experiential gestalt of causation as a possible underlying element in scientific misconceptions. He considers how students extend the primitive notion, learned in infancy of an agent that physically affects an object leads to a sense of "the nearer, the greater the effect." Applying this "causal expectation" in unexamined ways leads to difficulty in learning concepts that call for different, potentially more complex causal patterns. These examples, drawn from an increasingly rich research base on how students structure their causal expectations, illustrate how default patterns can complicate understanding of complex subject matter.

Causal Default Tendencies in Thinking about Evolution

This paper explores ten dimensions of causal thinking that people must engage in to reason about evolution. Each dimension is formulated in terms of a tension because typically the mindsets associated with each dimension are polar opposites and thinkers tend to be pulled to one side or the other, rather than dwelling easily in between. One side of the tension represents the default pattern that thinkers typically engage in. It represents ways that most people are comfortable conceptualizing cause and effect patterns. For example, in the case of centralized versus decentralized thinking, most people are fairly comfortable with centralized causal mechanisms. The other side of the tensions represents patterns of conceptualizing cause and effect that are unfamiliar to most people and may even be seen as a violation of how causes and effects should behave. For example, many people assume that a central control structure exists even when none is in play (Resnick, 1985). They find it difficult to visualize how an outcome can emerge as it does in decentralized causality.

We hypothesize that when thinkers know how to handle the side of the tension that is more unfamiliar, they are more likely to be able to dwell in between and see how both sides, or one or another side of a given dimension can be in play in a given instance. So if someone understands decentralized structures, he or she will be more likely (after an initial “honeymoon” of seeing many possibly decentralized structures) to recognize instances of both centralized and decentralized structure. However, if someone understands centralized structures, he or she will be no more likely to understand decentralized structures because of the conceptual leaps that they need to make in thinking about cause and effect.

Why might such default patterns exist? What makes certain ways of thinking about cause and effect come more “naturally” to us than others? Research on infants suggests that we may be "hard-wired" to expect certain causal contingencies such as priority, mechanism, and determinism (Bullock, Gelman, & Baillargeon, 1982). Determinism is the expectation that physical events are caused. Priority refers to the temporal ordering of causes and effects and an expectation that the causal relation is always unidirectional, with causes either preceding or coinciding with their effects. Mechanism refers to the assumption of the transfer of causal impetus, either directly or through a chain of events. Humean conceptions of causality fit with those above and lead to similar expectations. Hume’s categories include: 1) Antecedence or Temporal Precedence, "causes occur prior to or simultaneously with their effects"; 2) Covariation or Constant Conjunction- "a causal relation describes an invariable connection between events, so causes and their effects must systematically co-vary”; and 3) Contiguity in Place and Time- "causes must be contiguous in place and time or at least linked by an intervening chain of contiguous events" (Rosenthal & Rosnow, 1990). Certain patterns fit better with these causal contingencies than others. For instance, centralized causal control structures have a more obvious causal mechanism than decentralized forms. It is easier to “tell the causal story” of how one thing lead to another when the mechanism is a more obvious one.

The following paragraphs explain ten dimensions relating to causality that people must engage in when reasoning about evolution and explains how each relates. It describes the typical default patterns and analyzes why each tends to be most comfortable and where most thinkers tend to dwell. The members of the set of dimensions are not mutually exclusive which is why they seem best described as dimensions that have particular centers of gravity. We relate each dimension to an understanding of evolution and consider how each complicates attempts to understand evolution.

Dimension One: Linear vs. Radiating Causation

This dimension refers to the causal patterns of impact that people expect and subsequently what types of causal relationships they are sensitive to. Research suggests that people tend to expect efficient, simple, linear causal forms over more extended forms. This includes a tendency to seek out one to one correspondences rather than one to many or many to many correspondences.

In simple linear causal models, one thing typically makes another thing happen in a domino-like pattern of effects. There is temporal priority between cause and effect--causes always neatly precede effects–satisfying the expectation of priority as outlined by Bullock and colleagues (1982). There is usually one cause and one effect (Grotzer, 1993). Children may learn to expect this pattern from an early age. Andersson (1986) discusses the "experiential gestalt of causation" in which infants learn that there is an agent and it impacts an object. The relationship is one-to-one. Intensified efforts by the agent are believed to lead to an intensified impact on the object.

This dimension bears upon the mental models that people hold for the pattern of how evolution "progresses" as well as where it has "come from." According to Lord and Marino (1993) many people, including university students view evolution as "a slow linear movement from a single cell to human being" (pp. 353). The dimension pits a view of adaptation as linear and progressing neatly in tension against one that is radiating and includes the notion of lots of "throw-away" designs in nature's wastebin. Conceptualizing the "pattern" of evolution as radiating as opposed to linear helps in understanding at least two aspects of evolution: 1) The idea that natural selection works by generating and throwing away many designs. This relates to the inverted pyramid of diversity that S. J. Gould (1979) talks about; the great diversity of organisms that are on earth today are actually variations on a few limited themes, for example the ungulate species, or the insect. From the time of the Cambrian explosion there has been a general reduction in global diversity (although there is a phenomenal diversity today, it is based on a limited number of themes); 2) The second is that the general process that evolution follows is adaptive radiation or divergent evolution. Species with the most divergent evolution may experience a greater success in terms of adaptability. For instance, the great variety of bacteria suggests that bacteria may experience a long presence on the earth. Species such as humans or the horse are often presented as "classic" examples of evolution. It can be argued that these are actually the less successful species (in terms of diversity of extant species from that line of evolution). They could be seen as nearing the end of the line. Using these as "classic" examples in teaching evolution reinforces the notion that evolution is a linear process.

As mentioned above, the dimensions represent centers of gravity, not mutually exclusive categories. Understanding radiating causal patterns contributes to the following understandings and other dimensions. It helps students to conceptualize the "messiness" of the process and that lots of designs are "throw-aways" (and that successful designs cannot be predetermined by the creative process.) It helps students recognize the difficulty of prediction in determining what designs will be successful. It reinforces the importance of biodiversity. Finally, it helps students understand the dimension below on optimizing vs. satisficing because it contributes to a recognition that numerous designs can serve certain design purposes.

Dimension Two: Deterministic vs. Probabilistic Causation

This dimension addresses the tendency towards determinism and the difficulties learners have in dealing with uncertainty in general, and probability in relation to causality, in particular (e.g. Kalish, 1998). Students find it difficult enough to reason about probabilities (e.g. Tversky & Kahnneman, 19xx). Combine this with the challenge of how causes should behave and the difficulty becomes that much greater. Research shows that while five year olds are not accurate enough to detect instances where there is less that 100% covariation between causes and effects, slightly older children come to expect it (Sedlak & Kurtz, 1981). They use 100% covariation as evidence of a causal relationship (Grotzer & Bell, 1999; Kalish, 1998). According to Bullock and colleagues (1982) determinism is one of our key assumptions about the nature of causality. They demonstrated that even the youngest infants show puzzlement when a seemingly "uncaused" event occurs. Kalish (1998) found, while studying students' notions of disease transmission, that children do not recognize a mechanism as causal if it does not consistently lead to the specified outcome. Probabilistic causation is difficult for them to understand and this may contribute to risk-taking behaviors.

One of Darwin's main ideas in his theory of natural selection was that there must be variation within a population. Without this variation there is no potential for differential reproduction and therefore no potential for the selection of more successful organisms. Understanding the importance of variation lends one to see the importance of randomness in evolution. It is this randomness which leads to a type of order. Most people view randomness as destructive, but in some cases it actually helps make systems more orderly. The concept that greater variation can lead to greater stability is counterintuitive. This order is manifested in organisms being well adapted to their environments. So rather than viewing randomness as destructive it is important to view randomness as one of the creative forces in evolutionary systems. Understanding the probabilistic nature of evolution enables students to tap into their understanding of the creative forces of evolution.

Conceptualizing evolution as probabilistic recognizes the possibilities of combinations of gametes and also the importance of factors that introduce uncertainty, such as mutation. Understanding evolution involves reasoning about chance at many different levels. At the basic level of Mendelian genetics, at the level of chance produced from the generation of variation through genetic recombination and mutation, and at the level of environmental chance.

Even if an organism is well adapted to the environment it does not mean that it will be successful. Of course when you look at this from the population level this environmental randomness does not have as much of an effect. Still having a given set of characteristics is no guarantee of successful adaptation in an individual sense. There is a certain amount of randomness in which designs succeed over others. It is possible for more than one design to fill a niche but for one population to gain a hold on that niche over another for reasons that are completely due to chance. (For instance one population is wiped out due to the chance that they happen to be in the wrong place at the wrong time during a flood or a meteor strike.) Then perhaps due to their smaller numbers have a harder time competing with another population for the same niche.

Understanding evolution as probabilistic contributes to understanding that evolution is in fact a “random” set of individual events. It helps in recognizing the “messiness” of the process and the difficulty of prediction.

Dimension Three: Simple vs. Complex Interactions of Variables

This dimension relates to how people think about compound causal forms. When confronted with a set of interacting variables, it appears that people first assume a fairly simple set of interactions between them. They tend to assume simple, efficient causes and additive complex causes over other types of interactions such as multiplicative, diminishing, etc (Grotzer & Perkins, 2000).

To understand the complexity of evolution it is necessary to view it as a part of a larger system. The Modern Synthesis (understanding evolution through the lens of genetics) gave new insight into the workings of evolution. To understand the effect of natural selection on a gene pool is conceptually more difficult than understanding the natural selection of the individual within a gene pool. The Modern Synthesis enabled biologists to understand evolution at the gene pool level. While this understanding is vastly more complete it is also vastly more complex. This dimension addresses some of this conceptual complexity.

In attempting to understand the interaction between natural selection and the gene pool, one first assumes a fairly simple set of interactions. This simplistic view limits one’s understanding of population genetics and evolution. The Hardy-Weinberg principle states that there is a set of conditions that must be true for gene frequencies to remain constant over time in a gene pool. These conditions are impossible to create for an extended time in a natural population. However they offer a list of causal agents that effect gene pools. Natural selection is one of these causal agents. The complexity enters when one considers the interaction between natural selection and the other causal agents. Therefore one can find complex interactions between natural selection, genetic drift, gene flow, sexual selection and others. Understanding this complexity is a mark of deep understanding of evolution's place within a complex causal network. This understanding relates to understanding complex causal networks and mutual sufficient cause. Natural selection is only one of the sufficient causes in the list of Hardy-Weingberg conditions.

What follows is a specific example of this complex causal network. The Irish Elk is an example of a species where two types of selection acted upon the population and gene pool (In reality this is always the case, it is just that the Irish Elk significantly displayed the effect of two types of selection). The Irish elk was greatly affected by sexual selection. Sexual selection is a type of natural selection that is created when traits that directly effect an organism's potential for reproduction are selected for. These traits are often at odds with other selective pressures that focus on the survival of the organism (which in turn affects the reproductive fitness of the organisms). The trait that was selected for was the size of the Irish Elk's antlers. They were huge! Some of the antlers that were found were 6 feet across. It is thought that these huge antlers were a benefit in attracting a female mate. The larger the antlers the more "attractive" the male would be to the female. Therefore there is intense selection for larger and larger antlers. These large antlers could however be at odds to the survival. It is difficult to run away from prey if you are hauling huge antlers. Therefore a dynamic is created between classic natural selection and sexual selection. In this sense it seems that this is not a simple interaction but a complex interaction.

Understanding the complexity of evolution contributes to recognizing the difficulty of prediction and realizing that the likelihood of creating overly simplistic predictions is high.

Dimension Four: Intentional vs. Nonintentional

This dimension addresses the tendency to assign intentionality to controlling causal mechanisms. People tend to expect control that involves intention (and often a centralized form of intention, as discussed below–all working in concert towards the same goal.) However, sometimes the intention is an individual intention. Zimmerman (1987) found that 17% of a sample of high school biology teachers (n = 1000) believed that evolution involved the purposeful striving of an organism towards a higher, more complex form of life. Not surprisingly, Lord and Marino (1993) found that over 40% of students believed it, too. While natural selection assumes a similar rule (All individuals attempting to maximize their genetic success), it does not involve a purposeful, coordinated and well-intentioned set of goals. It has individual intention, but not group intention.

Beyond the distinction of individual or group intention, the way that intention plays out, is often not the way that students think of it. Individual intention might be involved in maximizing one's genetic success, but the options available for doing so tend to be limited. For instance, an organism can't evolve a longer neck because it wants one. However, students often engage in just this type of teleological thinking. One of the major pit-falls that students fall into is thinking of evolution as having a direction or a purpose. So organisms evolve longer necks because they "need" longer necks. Put another way, evolution is directed towards longer necks. It might be that evolution results in longer necks, but it is inaccurate to view this as the cause of the evolution.

In some respects, this tendency could be considered a subvariety of centralized vs. distributive control in part due to the nuance of the misconception. While assigning intention often accompanies centralized control it can also be an independent dimension of a larger misconception. One can mistakenly assign centralized control and then mistakenly also decide that it is governed by a purposeful monitor. Another argument for separating it out as an independent dimension, is that it is possible to come up with purposive systems that lack centralized control. On the other hand, hearts and livers are centralized but only purposeful in a weak sense. God or Gaia sets forth a strong purposive and centralized control mechanism.

To view evolution as a centrally controlled system can take form in two ways. It can take form as believing in a god who controls the evolution of species. Thinking about evolution with this direction, and purpose is a form of centralized control. This can be clarified by juxtaposing it to a decentralized system where the trends of evolution are emergent. The production of longer necks comes from many organisms acting out Darwin's natural selection that produces a complex system of the evolution of a species with longer necks.

Dimension Five: Centralized vs. Distributed Control

This tension addresses the tendency to expect centralized forms of control as opposed to distributed ones. It considers where learners look for causality in a system and what the nature of that causality is. Resnick and colleagues have shown that people tend to expect centralized control structures, for instance, a leader who determines an outcome (e.g. Resnick, 1996). According to Resnick (1996), "people tend to look for the causes, the reason, the driving force, the deciding factor. When people observe patterns and structures in the world (for example, the flocking patterns of birds or the foraging patterns of ants), they often assume centralized causes where none exist. And when people try to create patterns and structures in the world (for example, new organizations or new machines), they often impose centralized control where none is needed" (pp.)

Understanding evolution as a "random" set of individual events with emergent effects is cognitively very difficult. The causal agent is difficult to identify due to its distributed nature.

Thomas Aquinas's “Argument From Design” helps to illustrate why decentralized control is so difficult to grasp (Resnick, 1996). The story of this argument is the watch and the watchmaker. If you were walking in a field and came across a pocket watch lying on the ground you might pick it up and open it up to look at the mechanism. Once you had the watch open you would be amazed by the complexity of the gears, springs and mechanisms that make the watch work. With this amazement with the workings of the watch you would assume that this watch had been created by a watchmaker. You would not assume that it spontaneously came together there on the ground. With this analogy in mind the “Argument from Design” is that the complexity of life on Earth is so overwhelming that you would assume an omnipotent creator. Resnick (1996) talks about this tendency to see the complexity of life as centrally controlled. To see it as the product of decentralized control is challenging. However, most evolutionary biologists believe that this decentralized control is exactly what creates the complexity of life on Earth (refs).

Centralized patterns are created either by lead or seed according to Resnick. If they are created by lead, it assumes that there is a leader orchestrating the pattern. For example when looking at the ant colony one assumes that the Queen is controlling foraging behavior. If they are created by seed, some preexisting built-in inhomogeneity in the environment gave rise to the pattern much as a grain of sand gives rise to a pearl (Resnick p. 123).

We believe that understanding this decentralized nature of Darwin’s theory of natural selection is central to the understanding of evolution. If students understand decentralized systems it stands to reason that they would be more likely to understand the difference between teleological and non-teleological arguments for evolution. In this way understanding the decentralized control of evolutionary systems combats one of the major misconceptions about natural selection.

Dimension Six: Pattern vs. Serendipity

This refers to the human tendency to seek out and attach importance to patterns perhaps at the risk of missing the importance that serendipity plays in ultimately generating the descriptive patterns of evolution. It relates to the possibility of prediction because what is most likely to be filtered out as noise may be highly likely to generate significant pattern.

Totally random events that ultimately generate a pattern are part of the evolution process. In fact the fuel for natural selection is the generation of variation. A lot of variation ultimately has little or no impact on the species in the future (the patterns coming from white noise). Looking at a gene pool it is impossible to pick out the variation that will be selected for. Even knowing all the selective pressures (an impossibility) would not allow you to choose the genes in the gene pool that would have the most effect. This is because genes do not operate in isolation. They operate in individuals that operate in populations that operate in communities that operate in ecosystems comprising a very dynamic, unpredictable system.

This tension (perhaps unlike others) involves a balance of the two points. It also relates to the issue of purposeful vs. randomness- missing this can lead to teleological arguments, moving to valueless, non-purposeful, non-foresighted change.

Dimension Seven: Static vs. Dynamic

This dimension addresses the tendency to assume that evolution occurred within a static environment and to map changes with the assumption that we know the environmental pressures that led to certain adaptations. Evolution needs to be viewed as a mapping of adaptability to a dynamic and changing world. According to Levins and Lewontin (1985) “No species can ever be perfectly adapted because each is tracking a moving target.” Understanding evolution systemically involves recognizing that the earth is gradually undergoing changes (Cummings, Demastes, & Hafner, 1994).

It is certainly more difficult to assess outcomes against a dynamic backdrop than a static one. Evolution needs to be viewed as “a mapping of adaptability to a dynamic and changing world.” We tend to assume that we know the environmental pressures that led to certain adaptations. For instance, a cladogram (as is found in many Natural History Museums) helps to summarize adaptations that in hindsight appear to have enabled survival, but it may be that the particular adaptation was necessary but not sufficient for survival and that the co-occurrence of any one of some other set of adaptations with the particular adaptation together was necessary and sufficient. There may have been other animals around who “enabled” the survival of a certain animal and therefore their “adaptedness” is due not only to their design features but to environmental elements.

Dimension Eight: Individual vs. Population Reasoning

Students have a tendency to reason about effects on individuals rather than populations and to try to related evolution to “what it means for oneself.” According to Leach and colleagues (1992) this is particularly common in students up to the age of 13 years. Teachers may attempt to connect the lessons of evolution to students’ experiences in order to help them relate to the material and this can exacerbate this misconception.

This dimension involves difficulties in realizing the target of effects. Learners must move from a belief that individuals become more adapted over time to the belief that there is an increase in the proportion of adapted individuals over time. This is a critical leap for students to make in their understanding of natural selection. Unfortunately teachers at times confound thinking about evolution from the individual level by teaching about evolution out of context of population genetics. Many classes discuss population genetics but it is separated from the discussion of natural selection.

For instance, in the case of Darwin's finches, students often develop their understanding of natural selection focusing on the individual finches. However the individual finches are only important in their inclusion in the gene pool of finches that are under the influence of selective pressures. Students take the statement that finches developed larger beaks as individual finches "growing" larger beaks. This contradicts the theory of natural selection and ties in directly with teleological thinking about evolution. A more accurate conception of the statement that finches developed larger beaks is that finches as a population, or more accurately gene pool, existing over many generations increased its mean beak size. The rules that govern individuals are very different than the rules that govern populations. This is a difficult concept for students of biology. By understanding population genetics and population thinking the nature of the causal mechanisms of evolution develop more rich understanding of the mechanism of natural selection.

Dimension Nine: Predictive vs. Descriptive

This dimension refers to the expectation that causal mechanisms should lead to repeatable, predictable results. Even if one could hold all other variables constant, the same mechanisms would not lead to repeatable results because of the role that variation such as mutation play in the system. Whether descriptive causal patterns are repeatable such that they have value in a predictive sense is highly ambiguous. This dimension addresses the basic research question of evolutionary biology. The "experimental" foundation of evolutionary biology is based on putting together facts from the past. There are examples of natural selection that have been documented i.e. the Grant studies of the Galapagos Finches. However the vast majority of the ideas of evolution are based on piecing together evidence from the past. Generation of evolutionary models is a way in which evolutionary biologists attempt to make predictions. These models are based on the premises of the past. Yet at the same time, there is a recognition that rewinding the evolutionary taperecorder will lead to different outcomes. This dimension involves recognizing that we can describe what happens but that it is nearly impossible to model what will happen.

Dimension Ten: Optimizing vs. Satisficing

People tend to believe that in order to be successful a design must be "perfect" as opposed to good enough. However, many designs are "good enough" to be successful within the given environmental conditions. For instance, sharks constitute an effective design that is "good enough" to be successful. This tendency may relate to the common practice of teaching about humans as the logical end of an evolutionary path. Relating back to religious views of humans as “made in God’s image,” it may be more difficult to escape the view that designs must somehow be optimized. This beliefs ties into the notion of divergent evolution, the idea that there are many possible designs that could satisfy the particular environmental constraints of a given point in time and that there isn’t one perfect design that is necessary.

Methods

Overview:

Sixty two tenth grade students participated in the following design. All students were given a series of activities (not related to evolution) to assess their understanding of the ten causal dimensions. These tasks were computer based (including simulations in Star Logo (Resnick, 1985)) and paper and pencil based activities. Each task measured one or more causal dimension. These instruments were pilot-tested with a group of 60 students who did not participate in the formal study. A teacher- developed unit on evolution was taught. Students then participated in two assessments: 1) to reveal their depth of understanding of the process of evolution and; 2) to reveal their understanding of the causal dimensions as they are embedded in a understanding of evolution. In addition, students’ Preliminary Scholastic Achievement Test (PSAT) scores were collected to function as a measure of general achievement. The data was examined for correlations between understanding of the causal dimensions and understanding of evolution.

<u>Subjects</u>

The subjects were students in three biology classes at a private school near Washington D.C. There were approximately equivalent numbers of boys and girls. The students tend to come from upper SES families and therefore this might limit the generality of the results found here.

<u>Assessment Tasks:</u>

Evolution of Fashion Task:

It was decided to assess the causal dimensions using a topic that would be likely to draw out students’ best performance on the causal dimensions to see what they were capable of. We reasoned that students of adolescent age would be likely to feel comfortable reasoning about the evolution of fashion styles and that they would be likely to exhibit some of their most advanced reasoning on such a task. For instance, we considered that students would be more likely to mention non-centralized control structures for the kinds of fashion patterns that they observed even though they might also recognize centralized forces.

Students completed a set of multiple choice items with two questions for each causal dimension. The questions were given as statements ("If we could turn back time to the 1960s and then start the clock ticking again, fashion styles would end up exactly as they were the first time.") accompanied by an ordinal scale for students to register their level of agreement (from "I disagree a lot" to "I agree a lot"). The pair of questions were worded in opposite ways to allow us to check the reliability of the question such that students would answer high on one and low on the other. This was the case in the pilot data that we collected. The multiple choice data was scored by reversing the score for the negatively worded version of each question set and adding that to the score of the other question in the set, yielding a resulting score for each dimension.

We decided that it was important to offer other opportunities to show their reasoning because there is the possibility that the fashion task assessed their beliefs about the fashion world rather than their ability to reason along the various dimensions. For this reason, we included a number of other tasks designed to assess dimensions of reasoning that we believed may not come through as clearly on the evolution of fashion task.

Flocking Birds Tasks:

Students were given a task in which they were asked to explain reasons for the pattern of birds flocking as they saw it in 1) a videotape and 2)a StarLogo Program. We reasoned that it may be possible that students would give a more decentralized response to the computer program if they understand typical programming rules and a more centralized response to the videotape. Therefore, we administered both tasks and counterbalanced the administration across the different classes.

This data was scored by counting the number of centralized and decentralized statements made by each students and assigning an overall score. The following kinds of statements were scored as centralized, "the leader bird tells the rest of the flock where to go" and "the birds have a leader who flies first, then they all line up behind it and follow it." The following kinds of statements were scored as decentralized, "birds like to fly near other birds, so they all get their wing tips as close to another bird as they can" and "each bird gets behind another bird because it makes it easier to fly there and they end up in a pattern."

Two scorers scored all of the data and checked for reliability ($\underline{r} = .93$). Using a Pearson Product Moment Correlation. Differences were discussed and resolved until 100% agreement was reached.

Evolution of Species Tasks:

Following the curriculum unit, students were given an open-ended assessment of their understanding of the evolution of species. For instance, it included questions such as, “When they were first sold, insecticides were highly effective killing flies and mosquitos. Today some 20 years later, a much smaller proportion of these insects dies when sprayed. Explain why you think this might be true. How would this change occur?”

These assessments were scored using criteria developed in consultation with a professor of evolution at the collegiate level. For instance, for the question above, the students' answers were scored for whether they reflected a Darwinian or a Lamarckian model for evolution. For a Darwinian model, it was expected that students would identify that there was a “mutation” that existed BEFORE exposure to the insecticide that made the insects resistant to that insecticide. These individuals were “selected for” and reproduced more. (Students might also talk about increased fitness of the individuals with the mutation). Students would talk about the frequency of the gene increasing in the population over time. This is how the insects become resistant to the insecticide. For a Lamarckian explanation, students would be expected to talk about a NEED to become resistant to the insecticide. In this case, students typically reason that it was the exposure to the insecticide that caused the insects to become resistant. This is distinctly different than in the Darwinian explanation where the mutation existed before the exposure.

Students also completed a set of multiple choice items with two questions about evolution of species for each causal dimension. The questions were statements (“An organism can adapt if it tries hard. For instance, if giraffes need longer necks to reach the trees, they should try to stretch so that longer necks evolve.”) accompanied by an ordinal scale for students to register their level of agreement (from “I disagree a lot” to “I agree a lot”). Students were explicitly told that circling “I agree a lot.” means that they agree based on Darwin's theory of natural selection and that when they circle “I disagree a lot” it means that they disagree based on Darwin's theory of natural selection. The questions were not asking about their personal beliefs, rather about their understanding of the theory that they had been taught. The questions were paired such that there were two questions focused on each causal dimension and so that they were worded in opposite ways to allow us to check the reliability of the question such that students would answer high on one and low on the other. The multiple choice data was scored by reversing the score for the negatively worded version of each question set and adding that to the score of the other question in the set, yielding a resulting score for each dimension.

Results

Comparisons were made between students who demonstrated understanding of the embedded and unembedded dimensions and the scores they received on depth of understanding of evolution. Pairwise Pearson Product Moment Correlations were run to assess 1) how performance on various dimensions of the Fashion Multiple Choice Task correlated with performance on particular dimensions of the Evolution of Species Multiple Choice Task; 2) how performance on the Open-ended task correlated with performance on other tasks and 3) how performance on various dimensions correlated with one another. We also looked in depth at the various questions on the Open-ended Assessment of Evolution to assess how each correlated with understanding of the dimensions.

1. Did performance on certain dimensions of the Evolution of Fashion Multiple Choice Task correlate with performance on the same dimensions on the Evolution of Species Multiple Choice Task?

We were interested in whether student reasoning about the various dimensions of the Evolution of Fashion Multiple Choice Task would in any way predict their performance on the same dimensions of the Evolution of Species Multiple Choice Task. We hypothesized that the underlying causal reasoning might in some way contribute to success or difficulty on both tasks. There were strong correlations to be found on the following dimensions: Dimension Two-Deterministic vs. Probabilistic ($\underline{r} = .49$, $\underline{p} = .0007$); Dimension Nine- Prescriptive vs. Descriptive ($\underline{r} = .45$, $\underline{p} = .0077$); and Dimension Seven- Static vs. Dynamic ($\underline{r} = .37$, $\underline{p} = .0319$) reasoning. These results suggest that these might be promising areas to focus on in future intervention studies.

Correlations between the other dimensions were weak. The lack of correlation between the other dimensions may be interpreted in a number of ways. It is possible that the underlying reasoning about the various dimensions did not transfer between the fashion and the evolution tasks. Or perhaps students reason in ways that are situation-specific and the results should be interpreted to mean that it is not reasoning about the dimensions that creates difficulty, rather the situational variables govern students' responses. Another possibility is that students were able to reason in more sophisticated ways than the fashion task revealed, however, it did not come through because they were answering based upon their beliefs about the evolution of fashion and that these lean towards the less complex end of each tension.

2. Does performance on the open-ended assessments of the evolution of species correlate with performance on the multiple choice dimensions of the evolution of species?

Students' performance on the Open-ended Questions on Evolution (which does not focus specifically on the dimensions) was found to correlate positively ($\underline{r} = .44$, $\underline{p} = .002$) with how students score on the Evolution of Species Multiple Choice Task (which is designed to tap information about how students think about the dimensions.) This correlation is important because it suggests that performance on open-ended questions based upon criteria that teachers of evolution find important correlates positively with performance across the various dimensions. This task does not involve the issue of transfer of reasoning as does the fashion task. It affirms that reasoning well about the various dimensions as measured by the multiple choice task correlates with reasoning well about evolution more generally.

Could student performance on both the dimensions and the open-ended questions simply be explained by I.Q. or some equivalent measure? In order to examine this possibility, correlations between PSAT (as a stand in for I.Q.) and the evolution of species open-ended and multiple choice tasks were run. As one might expect, how students performed on the Open-ended Questions on Evolution correlates ($r = .30$, $p = .06$) with how students performed on their total PSAT scores. However, PSAT scores have a very low correlation with how students perform on the Evolution Multiple Choice Task which focuses on the dimensions ($r = .07$, p = ns). This suggests that the contribution to the Open-Ended Task scores made by the dimensions and that made by I.Q. or PSAT score are different ones. When one looks in greater detail, PSAT scores correlate at a fairly high level with some questions on the Open-ended Questions and not at all with others. PSAT scores correlate with some dimensions (Intentional vs. Non-intentional($r = .42$, $p = .05$) and Simple vs. Complex ($r = .15$, p = ns), and not at all with others.

Student scores on the Flocking Birds Task (focused on Centralized vs. Decentralized reasoning) did not reveal much in relation to student scores on the Open-ended Evolution Task. The correlation between student scores on the Flocking Birds Task and total score on the Open-Ended Questions about Evolution was at the level of $r = .24$, non-significant. The correlation between the Flocking Birds Task and Total PSAT was low ($r = .09$, ns).

3. How Independent do the Dimensions Appear to Be?

As suggested in the introduction above, one would expect substantial overlap between the various dimensions. They appear to overlap in some aspects and to support each other. The findings of the data analysis support this. A number of the dimensions correlate highly with one another as measured by either the Evolution of Fashion Task or the Evolution of Species Task:

Predictive vs. Descriptive and Pattern vs. Serendipity- $r = .56$, $p = .000$
Simple vs. Complex and Predictive vs. Descriptive- $r = .60$, $p = .000$
Intentional vs. Non-Intentional and Static vs. Dynamic - $r = .43$, $p = .001$
Centralized vs. Decentralized and Intentional vs. Non-Intentional- $r = .33$, $p = .01$

For each of these, a logic can be found to their correlation. For instance, if one leans more towards believing that evolutionary processes are predictive as opposed to descriptive in retrospect, they would likely believe in the power of pattern over serendipity in what eventually happens in an evolutionary sense. Similarly, they would be likely to view evolutionary processes as simple rather than complex. A centralized view of the world (as opposed to a decentralized one) is more likely to accompany an intentional view than an non-intentional view. A centralized and intentional view typically fits with a view of a purposeful deity who controls what happens.

4. How did Performance on Each Question on the Open-ended Assessment about Evolution Correlate with Particular Dimensions?

We then considered each question on the Open-ended Assessment in depth to consider whether it correlated with the particular dimensions assessed on the Evolution of Fashion or Evolution of Species Tasks (and with PSAT score). Significant correlations or correlations approaching significance were found on some of the questions. These are listed below.

“Question 1b: Imagine a population of insects that existed on a remote island who had never been exposed to insecticides. Given what you know from the above question [Question 1a], would the island population also become resistant to the insecticides after 20 years? Explain.”

Correlations on two dimensions were approaching significance; 1) Static vs. Dynamic on the Evolution MC Task (r = .26, p = .08) and; 2) Intentional vs. Non-intentional on the Evolution MC Task (r = .39, p = 06.) However, the Intentional versus Non-Intentional Dimension correlated with PSAT Score (r = .43) suggesting that the variance could be explained by general achievement factors. Total PSAT Score correlated with performance on this question at a level approaching significance (r = .30, p = .06).

“Question #3: Do you think humans are going through evolution right now? Please explain your answer.”

Not surprisingly, there was a correlation with the Static vs. Dynamic dimension on the Evolution MC Task (r = .36, p = .01). This makes sense in that students who perceived the environment as static would be more likely to interpret human evolution as static. The Intentional vs. Non-intentional on the Evolution MC Task was significant (r = .36, p = .01) but again, it correlated highly with PSAT score (r = .43) suggesting that it might be explained by general achievement factors. The Individual vs. Population dimension on the Evolution MC Task was significant (r = .42, p = .005) as might be predicted by students’ tendency to reason about human beings as individuals rather than anonymous members of a population. The Optimizing vs. Satisficing dimension on the Evolution MC Task was also approaching significance (r = .28, p

= 06) which might be expected given students' tendency to consider human beings as the "optimal design."

"Question #4: Which organisms do you think are more evolved, humans or insects? Explain your answer."
Significant correlations were found with the Linear vs. Radiating dimension on the Evolution MC Task (r = .38, p = .01) and with the Individual vs. Population on the Evolution MC Task (r = .24, p = .03). This might be explained by students tendency to perceive of humans as evolving along a linear path and as in question #3, difficulty in reasoning about humans as anonymous members of a population.

"Question #5: Of the following diagrams (linear, branching, radiating), which ones do you think accurately show how organisms evolve? Please explain your answer."

Performance on the Flocking Birds Task was approaching significance in correlating with performance on this question (r = .33, p = .06). The Simple vs. Complex dimension on the Evolution MC Task significantly correlated with this dimension (r = .32, p = 03), however, it also correlated with PSAT Score (r = .15). It makes a lot of sense that the Simple vs. Complex dimension would correlate with an answer on this question. One would expect that students who view the process as more complex would be more likely to view the task as branching or radiating than they would linear.

Discussion

We believe that the results here offer weak support for the possibility that the various dimensions are important in reasoning about evolution of species. A number of the dimensions have modest predictive ability for students' performance on the Open-Ended Evolution Task. A number of them appear to be independent from typical achievement or I.Q. type measures.

However, we also recognize that the results raise more questions than they answer. The correlations are modest and suggest a number of more specific questions to be explored. For instance, the issue of transfer needs to be considered. The fashion task was only predictive of performance on the evolution multiple choice (dimensions) task on a few measures. It is possible that the fashion task measured belief rather than ability. Another possibility is that it is difficult to transfer the types of reasoning outlined in the dimensions. Does the lack of correlation between the Evolution of Fashion Task and the Evolution of Species Task signal that there was not enough support for transfer to take place or that the types of causal reasoning we are seeking do not generalize across domains? These are important questions to address that speak to the efficacy of focusing on the causal dimensions outlined here.

While a correlational study is useful for offering insights into what might be going on an intervention study would lend clearer evidence for the hypothesis that these causal dimensions are a primary source of students' difficulty in understanding the evolution of species. The results here do suggest that perhaps such a study would focus on a smaller list of dimensions–those that appear to correlate the most with performance on the open-ended questions.

References

Andersson, B. (1986). The experiential gestalt of causation: A common core to pupils' preconceptions in science. *European Journal of Science Education*, 8(2), 155-171.

Brown, D. (1995). *Concrete focusing and refocusing : A cross-domain perspective on conceptual change in mechanics and electricity*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April 18-22.

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W.J. Friedman (Ed.), *The developmental psychology of time* (pp. 209-254). New York: Academic Press.

Cummings, C.L. & Demastes, S.S. & Hafner, M.S. (1994). Evolution: Biological Education's Under-Researched Unifying Theme". *Journal of Research in Science Teaching*, 31(5) 445-8.

diSessa, A.A.(1993). Toward an epistemology of physics. *Cognition and Instruction* 10(2 & 3), 105-226.

Gould, S.J. (1977). Ever since Darwin: Reflections in natural history. New York: WW Norton.

Kalish, C.W. (1998). Young children's prediction of illness: Failure to recognize probabilistic causation. *Developmental Psychology, 34*(5), 1046-1058.

Leach, J., Driver, R., Scott, P., & Wood-Robinson, C. (1992). Progression in conceptual understanding of ecological concepts by pupils aged 5-16, Centre for Studies in Science and Mathematics Education, University of Leeds.

Levins, R. & Lewontin, R. (1985). The dialectical biologist. Cambridge: Harvard University Press.

Lord, T. & Marino, S. (1993). How university students view the theory of evolution. *JCST*, May 1993, 353-357.

Resnick, M. (1996). Beyond the centralized mindset. Journal of the Learning Sciences, 5(1), 1-22.

Resnick, M. (1985). StarLogo. Available: www.media.mit.edu/~Starlogo/.

Zimmerman, M. (1987). The evolution-creation controversy: Opinions of Ohio high school biology teachers. *Ohio Journal of Science, 87*(4); 115-125.