

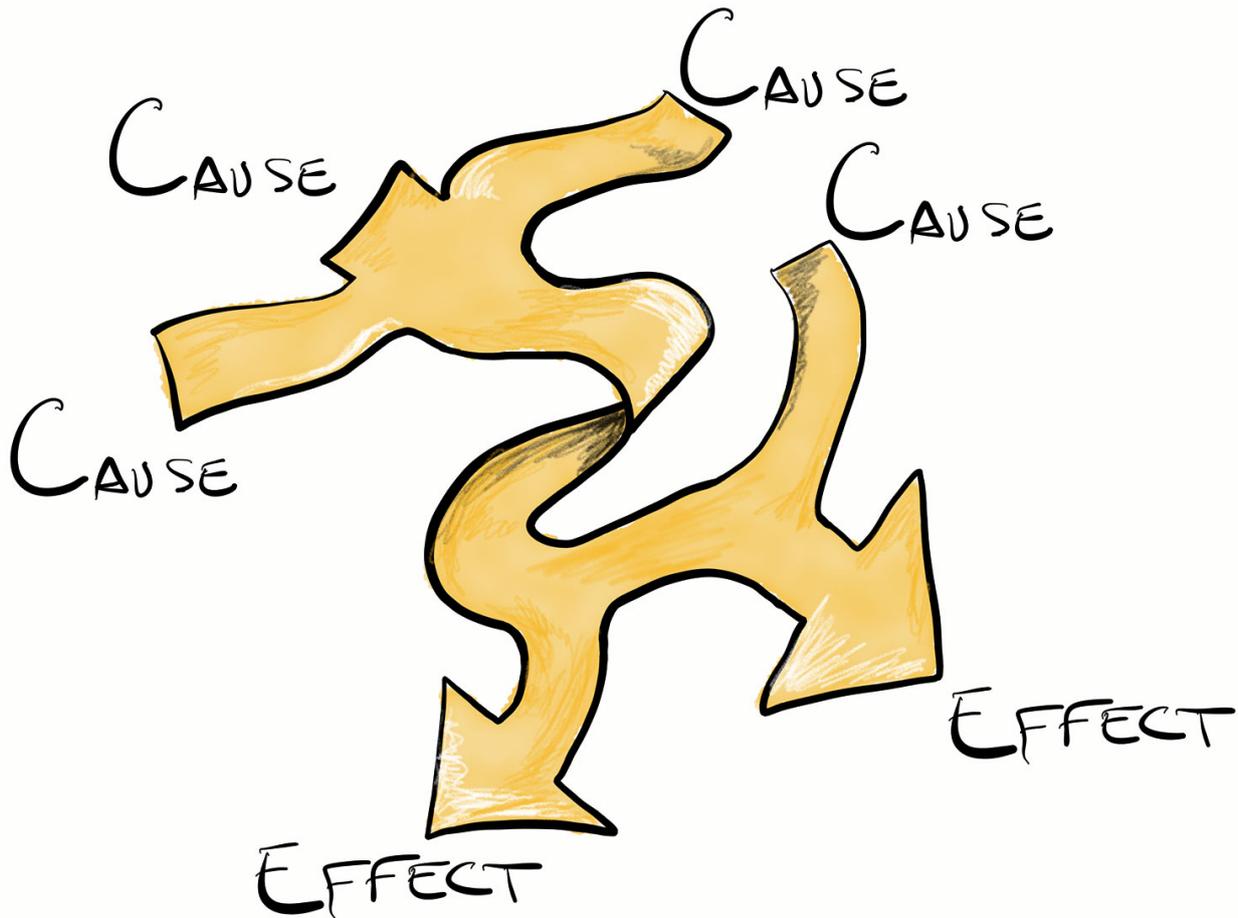


Solution: ‘The Slippery Math of Causation’

The all-too-intuitive picture of a straight arrow going from cause to effect is far too simplistic to describe the real world.

By Pradeep Mutalik

[Our latest *Insights* puzzle](#) attempted to model multifactorial causation with problems that involved three causal factors whose different types of interactions either produced or did not produce an effect. The goal was to challenge the all-too-intuitive picture of a straight arrow going from cause to effect as far too simplistic to describe the real world. In fact, a recent *Quanta* article by Veronique Greenwood describes the omnigenic model of complex traits, with the startling self-explanatory title “[Theory Suggests That All Genes Affect Every Complex Trait](#).” The thinking that inspired our puzzle and is manifest in the omnigenic theory was captured visually by [Paul Laurienti](#). In the diagram below, think of the genes as the causes and the complex traits as the effects, and you will see that it fits the new theory perfectly. Thank you, Paul!



Paul Laurienti

Causation is indeed a complex and slippery phenomenon, and readers contributed a wealth of interesting perspectives on it in their comments. I refer to and discuss reader comments extensively below. But first, let us look at the puzzles themselves.

Problem 1

Consider a scenario where there are three causative factors, a , b and c , which are real variables that take values between 0 and 2. The three factors interact together to determine the value of a hidden factor, d . If the value of d is within a certain window, then a particular event occurs (Y). If not, the event fails to occur (N).

[puzzles]

Consider three models of causation: i) a simple linear interaction between a , b and c (where the value of d is the sum of the three after each is multiplied by its own nonzero constant factor); ii) a “tennis serve” model where a , b and c are like the height, vertical and lateral angle of the hit and d is the position where the ball lands, which has to be in the service court; and iii) a “genetic model” where a , b and c are gene products, two of which interact multiplicatively to form an intermediate product that interacts linearly with the third gene product to determine the final concentration of d . The size of the window for d to result in the target event can be set

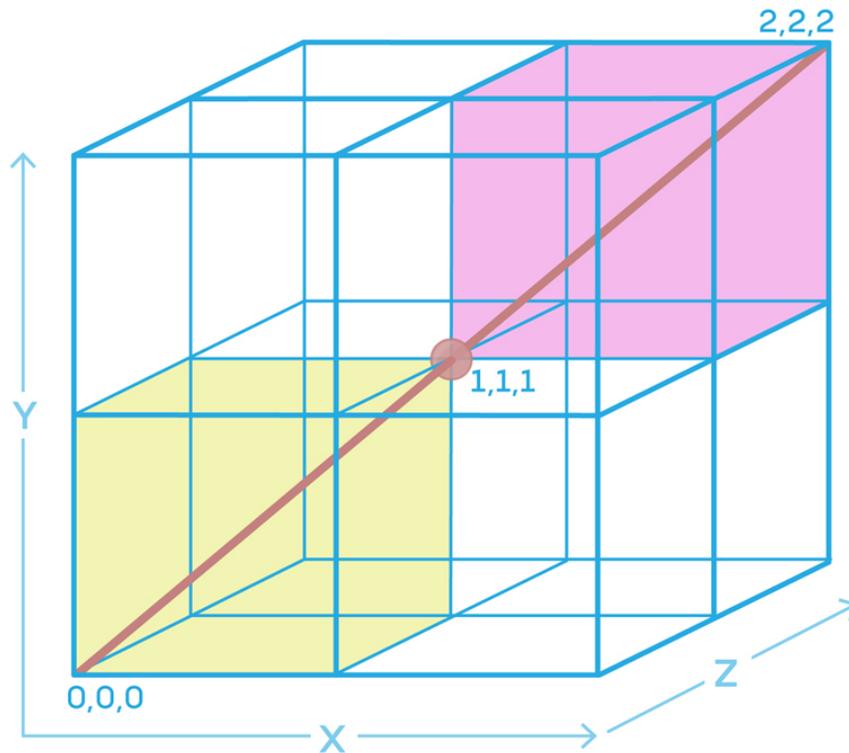
arbitrarily, but must be less than one-twentieth of the total range of values that d can take as a , b and c vary between their extreme values.

Which of the three models described can allow the target event to occur only when a , b and c are all greater than 1 or are all less than or equal to 1, but in no other circumstance? Can you think of a way that a , b and c can interact that will naturally give rise to this result?

The required result is shown in the table below. A “Y” in a specific cell means that the target event can occur somewhere within the sub-ranges of a , b and c values specified in the particular cell; an “N” means that it cannot.

	$0 < b \leq 1$ $0 < c \leq 1$	$0 < b \leq 1$ $1 < c < 2$	$1 < b < 2$ $0 < c \leq 1$	$1 < b < 2$ $1 < c < 2$
$0 < a \leq 1$	Y	N	N	N
$1 < a < 2$	N	N	N	Y

The answer is that none of the models proposed can produce the desired result. The key insight is topological. The above 8-cell table can be visualized in three dimensions as the eight “cubies” of a $2 \times 2 \times 2$ cube as [Mark Pearson](#) and [Zach Wissner-Gross](#) realized. You can visualize this cube as sitting at the origin $(0,0,0)$ of the xyz-coordinate system and extending to the point $(2,2,2)$. The x, y and z coordinates correspond to the values of the causative factors a , b and c values respectively, which assign a specific value for d to every point on the cube. The “Y” cells specified above then correspond to the cubies or octants that are located at the bottom left front and the top right back of the cube. These two cubies touch only at a single point $(1,1,1)$ as seen in the figure below. For the linear model, as [Wissner-Gross](#) put it, “the locus of points within some target window is a ‘3-D sandwich,’ a stack of planes. There is no way for such a locus to exist only in opposing octants of a cube. One can consider the line segment connecting the two solutions in opposing octants. That segment will either pass through another octant or the point $(1,1,1)$, in which case we can find another solution arbitrarily close to $(1,1,1)$ that lies in another octant. No matter what, the linear model cannot produce the desired table.”



Lucy Reading-Ikkanda/Quanta Magazine

Why does the target window have to be a 3-D sandwich? Because, as a consequence of the linear model, the target window must have length, breadth and height, and cannot pass through a single point. To understand this, consider the linear problem in two dimensions, where $d = px + qy$. Imagine a 2×2 square going from 0,0 to 2,2 on the xy-plane. Let $p = q = 1$, so that $d = x + y$, just the sum of the point's x and y coordinates. Every point on this square thus has a value for d , which is 0 at lower left, 1 at upper left, 1 again at lower right and 2 at upper right. In fact, all points on the diagonal from upper left to lower right have the value 1. Every subrange of d , such as a target window of 0.95 to 1.05, will be a 2-D sandwich between lines parallel to, and on either side of, the above diagonal.

The same applies in three dimensions. In our three-variable linear model, every point in the cube has a value $d = pa + qb + rc$, where p , q and r are constants. The fact that the three variables a , b and c extend from 0 to 2 in three dimensions means that no matter how much we try to flatten our 3-D sandwich, it always retains some finite thickness. As a concrete example, let $p = q = 0.995$. In this case the sum of pa and qb can take on values that range from 0 to 3.98 and at $a = 1$ and $b = 1$, the sum is 1.99. To flatten the sandwich, we set c 's multiplier r as small as possible. To make d exactly 2 at (1,1,1), r needs to be 0.01. We can set the target range of d to be between 1.9 to 2.1, which covers one-twentieth of d 's range. But now look what happens around the (1,1,1) point.

	$0 < b \leq 1$ $0 < c \leq 1$ $b = c = 1$	$0 < b \leq 1$ $1 < c < 2$ $b = 1,$ $c = 1.01$	$1 < b < 2$ $0 < c \leq 1$ $b = 1.01,$ $c = 1$	$1 < b < 2$ $1 < c < 2$ $b = c = 1.01$
--	---	---	---	--

$0 < a \leq 1$ $a = 1$	Y $d = 2$	Y $d = 2.00$	Y $d = 2.01$	Y $d = 2.01$
$1 < a < 2$ $a = 1.01$	Y $d = 2.01$	Y $d = 2.01$	Y $d = 2.02$	Y $d = 2.02$

As you can see, there are now Ys in all the cells! The sandwich, though flattish, extends over values of the variables in all eight cells.

A similar analysis can be applied to the other models as well, as [Wissner-Gross](#) showed.

What kind of model works? As [Wissner-Gross](#) correctly states, “An example of a model that can produce the desired table is

$$d = (a^2 + b^2 + c^2) \times [(a-2)^2 + (b-2)^2 + (c-2)^2]$$

“Namely, it is the product of the squared distances from (0,0,0) and (2,2,2). The target window can be set arbitrarily close to zero, and only points very close to those opposing vertices will be solutions.”

[Younes Rabii](#) also presented a well-reasoned analysis and gave a clever answer that used “max” and “min” functions. This could work, but using such functions goes against the spirit of the problem because it allows d to be computed with no contribution at all from one of the three variables.

Problem 2

What are the maximum number of cells in each of these models that can show a “Y”?

All the cells can have Ys as demonstrated above. As [Mark Pearson](#) explains, this is because of topological continuity. If two points lie in cubies that touch only at a point and have a d value in the target range, there will be very many paths joining these two points consisting entirely of intermediate points that also have d values within the target range. Many of these points must lie in surrounding cubies.

Problem 3

Of the 256 possible Y-N patterns that these tables can contain, which model can achieve the most? Can it achieve all?

No one attempted this problem. It is an interesting problem in its own right, so I will leave it open for readers to solve, with a hint. As we saw, our models cannot achieve a pattern with Ys exclusively limited to two opposite-corner cubies that only share a single point. On the other hand, many other patterns, such as Ys in all the cubies, can be easily achieved. But can linear models achieve a pattern that contains Ys exclusively limited to two cubies that only share a single edge? The question to ask is: Can a three-dimensional object lie partly in one of the cubies and partly in the other without passing through any other cubies?

What these problems show is that just three simply interacting factors (causes) can produce perhaps 100 different patterns of yes/no events (effects) of eight different types. Since the causes are mathematically similar variables, in a “fair” mathematical theory of causation, all of them should share equal responsibility for the outcomes. But in comparable real-world cases, we do not intuitively use mathematical equivalence to assign equal responsibility.

Consider this interesting observation from Mark Pearson: “Talking about causality always bothered

me reading reporting on politics. Consider the headline from last year: McCain Casts Crucial Vote to Kill ‘Skinny’ ObamaCare Repeal. When a bill fails 49-51 (note: this didn’t happen in this case), you can’t really pick out one person and say it was their vote that caused the failure. There are another 50 people who also voted no, and all their votes were necessary.” Yes, mathematically that is true, but in this case we pick the vote of the person who voted against convention, or against expectation, as the person who caused the failure. This shows that causality is an intensely practical concept that belies mathematical fairness. As I mentioned in the puzzle column, we intuitively single out as causes those factors that can be controlled or manipulated most easily, presumably as an evolutionary adaptation, allowing us to focus on changing those factors and thus causing a different result if similar situations should arise in the future.

I remember coming across this concept years ago in Isaac Asimov’s *The End of Eternity*, generally considered his best work. In it, a group of humans called “The Eternals” change history by going back into the past and causing the minimum necessary change (MNC) that would give rise to the maximum desired effect such as preventing a war. Thus a hypothetical MNC required to avoid President Kennedy’s assassination might have been something like disabling the car of the friend Lee Harvey Oswald rode to work with or nailing shut the doorway to the perch at the Texas School Book Depository from which he fired his rifle.

This example illustrates that another way we identify something as a cause is to reason counterfactually: to imagine what would have happened if a particular factor were removed or disabled. It’s no wonder that the fourth pillar of Judea Pearl’s “Eight Pillars of Causal Wisdom,” which he discusses in [his new book](#), is called “the [algorithmization of counterfactuals](#).” Counterfactual thinking helps in assigning the appropriate responsibility (or blame) to one factor among many and identifying the one that gives the most “bang for the buck.” Two such systems, discussed in a conversation in the comments between [Mark Pearson](#) and [Laurence Cox](#), are the Nobel Prize-winning [Shapley Values](#) and the [Chockler-Halpern](#) approach. These approaches try to balance mathematical fairness with practical counterfactual thinking — with results that work well in certain circumstances, but not in others. Other approaches that originated in industrial quality control and focus on the practical end of the spectrum are the [5 Whys](#) and Ishikawa’s [Fishbone Diagrams](#). As we can see, causation is clearly a hybrid beast — a true chimera — created by the marriage between mathematical fairness and practical problem solving.

It was interesting to see feedback from a couple of readers who actually knew about wildfire risk analysis. [Don Parsley](#), thank you for your insights about this topic, and I do want to assure you that my intention was not to discount clearly identifiable practical causes but to encourage discussion about the complex multifactorial aspects of causality. [SDY](#), I apologize for calling [flying hot wood embers](#) sparks; but then, you did call embers ambers. And [Suzanne Grady](#), thank you for drawing attention to the [patterns of repetitive causation](#) and all the areas where analysis of complex causation can provide nuanced answers.

With all these wonderful comments, it was difficult to pick a winner for the *Quanta Magazine* T-shirt. I loved Paul Laurienti’s diagram of multipronged causality. Mark Pearson came very close to winning, for his many contributions and discussions, which have been referenced extensively above. The winner of the T-shirt, by a nose, is [Zach Wissner-Gross](#) for his clear exposition of problem 1. Congratulations!

See you in August for new *Insights*.

Correction July 2, 2018: The second chart under problem 1 originally displayed both Ys and Ns. It has been corrected to show all Ys.