8

# SIMPLICITY AS FALL-OUT

If an indefinite number of hypotheses fit our data, we need some way to select which one to accept. Merely fitting the data is not sufficient; the additional criterion philosophers put forth most prominently is simplicity: we should believe the simplest of the hypotheses (we can formulate) that fit (and would explain) our data.

Why believe the simplest? Convenience in manipulating, remembering, transferring, and teaching the theory or hypothesis would lead us, perhaps, to utilize the simplest, but why *believe* it? Why think that, of the otherwise equally satisfactory explanatory theories, the simplest is most likely to be true? If simplicity is "relative to the texture of a conceptual scheme,"[1] to the kind of "graph paper" on which we plot the world, then to show this connection between simplicity and truth would involve the unpromising task of establishing our conceptual scheme or coordinate system as privileged.

The question about simplicity is not merely a skeptical conumdrum. It is difficult to think of *any* reasonable explanation for why including such a simplicity maxim should help (and should have helped) make the institution of science successful or more likely to arrive at the truth. To explain how a simplicity maxim contributes to the success of science by saying that the world is simple says little, if indeed it says anything other than that there is some simple true theory of the world. (Of *all* of it? And why think any particular simple theory, which explains only some things, will be a component of the overall simplest theory?) So the maxim of simplicity and its connection with truth puzzles us.

W. V. Quine and Joseph Ullian[2] find that natural selection "offers a causal connection between subjective simplicity and objective truth . . . Innate subjective standards of simplicity that make people prefer some hypotheses to others will have survival value insofar as they favor successful prediction. Those who predict best are likeliest to survive and reproduce their kind . . . and so their innate standards of simplicity are handed down" (p. 47). But maxims of simplicity direct us to choose the simplest of hypotheses in areas and at levels in physics, chemistry, cosmology, molecular biology, none of which were believed or chosen by our ancestors. What explains why their subjective standards of simplicity, which have stood the evolutionary test for hypotheses about middle-sized macro-phenomena, also connect with truths about cosmology or micro-phenomena? (Or is there no "connection between subjective simplicity and objective truth" in these areas?) Perhaps the fact that nature is uniform, here interpreted as meaning that the same type of simplicity obtains throughout. *If* one could turn this into a precise statement about *the world,* the requisite sort of explanatory link between our standards of simplicity and truth would be forged. Quine and Ullian further note that "such standards will also change in the light of experience, becoming still better adapted to the growing body of science in the course of the individual's lifetime. (But these improvements do not get handed down genetically)" (p. 47). However, they do get handed down in the education of scientists. If scientists' standards of simplicity change to fit the theories they develop and accept, then these theories will seem to them and to later generations, who judge by these changed standards, simple.

In "On Simple Theories of a Complex World," Quine artfully sidestepped the issue of simplicity's connection with truth by connecting it not with truth but with confirmation. A preference for simplicity is, so to speak, an *artifact* of our procedures of confirmation, so that the simpler of two theories is more likely to be confirmed and to hold up under the way we investigate theories. The question Quine raises, what the artifacts of our procedures are, is of extreme importance. However, to raise the question of whether something is an artifact of our procedures of confirmation is to raise the question of whether, in that respect, those procedures should determine our beliefs. If people who learn science (strangely) got headaches when they considered a hypothesis containing three words beginning with the letter *d*, then no such hypothesis could get confirmed. That would be an artifact of the man-scientific-procedures combination, but surely no reason to disbelieve such a hypothesis. Thus, questions of the justification, and not merely the causes, of our beliefs would be raised. Put somewhat differently, noticing the artifact and knowing the artifactual causal story might lead us to change our (artifactual) beliefs.[3]

Simplicity is usually introduced as an additional factor to help decide among the different hypotheses that fit (or would explain) the data $D_1$ in hand at time $t_1$. Another way to decide would be to gather more data. But this seems not to solve the current problem, for however much new data $D_2$ we gather, and however many old hypotheses $D_2$ eliminates, there will still be an indefinite number of hypotheses that fit $D_1 + D_2$. And how are we to choose among them?

The feeling that no progress has been made by collecting new data—to be sure some old hypotheses are refuted—should set us to wondering whether the situation is being misconceived. Instead of thinking of a temporal cross-section of hypotheses and data, consider the following very simple model of an ongoing process.[4] At time $t_1$ there are data $D_1$ and already formulated hypotheses $H_1 \ldots H_m$ which fit $D_1$. New data $D_2$ are gathered to *select among* these hypotheses. If only one hypothesis survives, believe it (tentatively), and go on to gather new data to test it against other hypotheses you now think up which fit $D_1 + D_2$.

"But how can it matter that a particular hypothesis was actually thought of earlier and tested? If another hypothesis $H_{23}$ which we have just now thought of and which also fits $D_1 + D_2$ had been thought of earlier, then it would fit all the data we now have. Surely which hypothesis the data support cannot depend upon which hypothesis we happened to be thinking about as we gathered the data." If $H_{23}$ had been thought of earlier it would fit $D_1 + D_2$ now, but we *don't* know that it would fit all the data we would have now, for if $H_{23}$ had been thought of earlier, we would have gathered *different* data. Our process of data gathering is selective. Given hypotheses $H_i : y = f_i(x)$ we gather data at values of $x$ where the hypotheses differ, and at a sufficient number of points $x$ so as to leave only one of the original hypotheses surviving. If $H_{23}$ had been one of that original number, we would have gathered data at some additional points $x$.

Granted that the counterfactual that purports to make the sequential procedure look ridiculous is *not* known to be true, still the question arises of why we should (tentatively) believe the surviving hypothesis rather than another that fits those data just as well. The answer depends on the way the underdetermination of theory appears to make it difficult to (strongly) support a hypothesis. Hypothesis $H$ fits data $D$. But an infinite number do. We expect many to, and this is one of those which do. Whatever the data we found, we could dream up many hypotheses afterwards to fit them, and so with $D$ we dreamed up $H$; is this any reason to believe $H$?

Suppose someone says he will show you that he is a very skilled archer. If an arrow is shot at random at a large target, the probability of its hitting a particular point is *very* small (has measure zero). And that's true for each point. "So," he says, "if I shoot and succeed in getting the arrow to a particular point, it will show I am very skilled." He shoots, the arrow lands, and he says, "See how skilled I am; the probability that the arrow would have landed precisely there at random was minuscule." "Hold on," we say, "the arrow had to land *somewhere,* and it landed there. Wherever it landed would have been an unlikely place. We'll believe you're a skilled archer if you *first specify* the minuscule area your arrow is going to hit, and *then* succeed in hitting it." Similarly the data are going to land somewhere. That they land in a particular place is a reason for thinking a specification of that place is significant only if *that* specification was offered beforehand.

Suppose the archer shot at random at a wall, and then went to where the arrow had landed, drew a circle around it, and said "Bull's-eye!" Suppose when the data land somewhere, we draw a curve through it and say "Bull's-eye." A bull's-eye is when the data hit the hypothesis, not the other way around. Since bull's-eyes give support, to believe the best supported hypothesis is (some complications aside) to adopt the method of tenacity.[5]

Now to connect simplicity with the sequential selective data gathering and testing procedure. We (tend to) think of the simpler hypotheses, relative to *our* conceptual scheme. So the simpler hypotheses get into the fray early. Suppose only one hypothesis $H_S$ from the initial batch of hypotheses we formulated survives our gathering of data $D_1$ and, in accordance with the procedure, we tentatively believe it. Either:

(1) The initial hypotheses[6] were the simplest possible which would fit $D_1$; no other possible hypothesis that fits $D_1$ is simpler than any one of these. With all but one of the original hypotheses eliminated

by $D_2$, $H_S$ is the simplest possible uneliminated hypothesis that fits $D_1$. Since whatever fits the large data set $D_1 + D_2$ must also fit $D_1$, $H_S$ is the simplest possible hypothesis that fits $D_1 + D_2$. So we *already* are (tentatively) believing the simplest hypothesis compatible with our data.

Or (2) we did not originally formulate the simplest hypotheses that fit $D_1$, and there are hypotheses simpler than $H_S$ compatible with $D_1 + D_2$. We formulate further hypotheses, and enter the next stage, gathering the (selective) data $D_3$, and we (tentatively) believe the surviving hypothesis. Either it's now like case 1 above: our formulated hypotheses were the simplest that fit $D_1 + D_2$ (that is, their subset bounded by this stage's survivor contained the simplest), and so we're now believing the simplest hypothesis compatible with all of our data, or else (loop) case 2 again.

This sequential testing procedure fails to get us believing the simplest hypothesis (of those we can formulate) compatible with our data only if there is an infinite descending sequence of more and more simple hypotheses, and we start up on it and move slowly down. If we sequentially test hypotheses that we've thought of, tentatively believing the survivor, and if we do (tend to) think of simple hypotheses (judged relative to our conceptual scheme), then the result will soon be that we are believing the simplest hypothesis (of those we can formulate) which is compatible with all the data we have. (Even simpler hypotheses, of course, may have been eliminated earlier.)

We need no rule or maxim that bids us to believe the simplest hypothesis compatible with our data. It's just a consequence of the operation of *that* sequential procedure that we *will* end up doing so. And the fact that simplicity seems relative to our own conceptual scheme, to our background concepts, coordinate system, type of conceptual graph paper, etc.—a fact that blocks attempts to connect simplicity with truth—fits naturally into *this* story. The more simplicity fits however we happen to tend to think, the more likely it is that we *will* think that way, and early. To *guarantee* we'll end up believing the simplest hypothesis compatible with the data, it would have to be assumed, not merely that we tend to think up simple hypotheses, but that by some time we have thought up *all* the simplest. This assumption is too strong, but then again science provides no guarantee of convergence to the simplest. It is no defect for our theory of simplicity to explain no more than the facts.

Simplicity is not called in as an additional criterion for hypothesis selection which is made necessary by underdetermination of theories. Rather, underdetermination of theories makes necessary, in order to have hypotheses supported, a sequential process of data gathering and testing, and out of that process drops simplicity (as judged by us).[7] But what is the explanation of the fact that we do find hypotheses that fit the data at all, and that not all those we think up are always eliminated at the very next stage? Scientists choose to work on problems and in areas where they think they can get results, and leave intractable problems aside. It would need explaining if they were never able to succeed!

Our argument about how simplicity precipitates out of the sequential process would apply to any situation in which:

(1) Instances of some type of thing *T* are generated.
(2) The instances can be ordered along some dimension *D*.
(3) Those *T*'s which are generated early tend to cluster around and eventually exhaust one end of the dimension; they are *D*-er than any *T*'s not yet generated.
(4) At each stage, *T*'s that fail to satisfy some criterion *C* are eliminated.
(5) The ordering of the *T*'s along the dimension *D* does not depend upon their relation to the criterion *C*.

Since these conditions do not focus especially upon simplicity, any process with a dimension *D* that satisfies 1 to 5 will eventuate in the survival of the *D*-est.[8]

The theory we have presented yields simplicity as fall-out from the process, but it gives simplicity, qua simplicity, no role as input, and so does not account for the way we utilize simplicity, seek it, and consciously favor it. For example, the sequential process may involve rejecting previous data on grounds of simplicity, if there is a relatively simple hypothesis that fits $D_2$ and most of $D_1$, and only very complicated hypotheses fit $D_2$ and *all* of $D_1$. Some of $D_1$ may be rejected in order to accept the simpler hypothesis. Or suppose we came upon beings on another planet with a different simplicity ordering and, hence, with a different history of hypothesis testing. They had never tried out what (to us) is the simplest hypothesis about the operation of their environment. Would we be impressed by the history of one of their hypotheses? Before we tested theirs against the simplest one of ours that fit their data, which would we believe? If ours, even though it hadn't yet been tested by them or us, that also would show that there's something more to simplicity than merely the fall-out result of the sequential procedure.

Our story about simplicity as fall-out can be elaborated to give simplicity some role as input to the process as well. Once we notice that we are believing the simplest hypothesis that fits the data, we can come to pursue a policy of doing so. The fall-out provides the basis of an induction. Here Quine and Ullian's natural-selection argument serves. We find ourselves having simple beliefs about ordinary middle-sized natural macro-objects and situations because those (who were contemporaries of our ancestors) to whom these truths seemed complicated left no descendants among our contemporaries. Also, since the process of sequential testing and tentative acceptance leads us to end up believing the simplest hypothesis compatible with our data, a retrospective look will find that simplicity is successful, that the successful are simple. Even more so because, as Quine and Ullian note, our culturally transmitted standards of simplicity tend to change so as to fit more neatly what we actually have ended up accepting.[9]

Since the past exhibits a correlation between the simplicity and the success of a hypothesis, a modest induction—to be sure, a simple one, but that's how we tend to think—leads us to conclude that these do go together and, hence, to rely upon simplicity. There are various ways to imagine the induction; one is as an inference to the best explanation.[10] We start with an observed connection between simplicity and acceptance by scientific procedures, and we explain this by positing connections between each of these and truth: we posit that acceptance of hypotheses by scientific procedures is correlated with their truth (there is independent reason to believe this, since scientific procedures eliminate *false* hypotheses that conflict with the data) and that the simplicity of hypotheses is connected with their truth (that the simpler of competing hypotheses is more likely to be true). These last two connections, if sufficiently tight, would imply and explain the observed correlation. One, therefore, might be led, in order to explain the observed connection, to infer a real connection of simplicity with truth. This account is admittedly rough, but fortunately we require no precise and rigorous inference for a plausible account of how one might actually arrive at a trust in simplicity. It might appear, however, that this account is vitiated by circularity. For the conclusion that simplicity and truth really are connected is inferred in order to explain the observed connection of simplicity with acceptance by scientific procedures. Yet there are other more complicated explanations of this observed connection. So in making the inference to the connection of simplicity with truth (and not some inference to a more complicated explanation) isn't a simplicity maxim already being used? A simple inference was made, but simplicity was not consciously pursued. Since our purpose is not to *justify* simplicity but to explain how it might come to be consciously pursued as a goal, we legitimately may place the wisdom of pursuing simplicity as the conclusion of an inference which instances simplicity without itself pursuing it.

Starting only with simplicity as fall-out, we end up trusting in simplicity and using it in the process. Seeking simplicity affects the sequential procedure in two ways: first, we try to think up especially simple hypotheses and, second, we consciously use simplicity as a criterion in selecting among hypotheses. Thereby the sequential process even more effectively yields simplicity.[11]

This lovely picture, unfortunately, is marred by the fact that our fall-out explanation of simplicity, if correct, undercuts the induction that leads to trusting in simplicity. If the fall-out tale is the best explanation of the observed correlation between success and simplicity, then the inductive extrapolation that leads to trust in simplicity is blocked; that is, it is unreasonable. But still, it may have occurred (and *not* unreasonably then, since no one then had the fall-out explanation), and that induction may be the actual explanation of our current trust in simplicity.

Does and should accepting all this undermine trust in simplicity? It would be pleasant to bring things full circle by having the trust tenaciously hang on. After all, the hypothesis that there is a real connection between the simplicity and the success of a hypothesis entered the field before my explanation did; that hypothesis fit past data, and new data were gathered which the hypothesis also fit, and so that hypothesis came to be accepted. In accordance with the method of tenacity, that hypothesis and the accompanying trust in simplicity will and should continue; that hypothesis should not be displaced merely because some new hypothesis has been thought up which also fits the data.

But even if we granted this application of the method of tenacity, our confidence in simplicity might well be short-lived; for the next step, now that another hypothesis has entered the field, would be to use the hypothesis to generate different predictions, to test them selectively, to discover which survives, and to end up believing that one.

However, I find myself *already* believing the explanation I have offered. This might be because the two explanations, mine and the one positing a real connection between simplicity and truth, do not fit exactly the same currently available data, because mine already explains more. But another view of the matter is more fun.

My explanation involving fall-out and the induction based upon the fall-out is, I believe, simple, elegant, forceful, and lovely. More so, surely, than the reigning hypothesis of a real connection between simplicity and truth. The simplest and most elegant hypothesis is that there is no real connection between simplicity (and elegance) and truth. Now if I accept this hypothesis as true (partly) *because* it is so simple and elegant, and the proponent of a real connection between simplicity (and elegance) on the one hand and truth on the other rejects this hypothesis as true because it denies any such connection, with whose petard is each of us hoist?