The Problem of Free Will in Program Evaluation

Michael Scriven

A group of hard-nosed scientists who have been studying the major commercial weight-loss programs recently reported their disappointment that the proprietors of these programs refuse to release data on attrition. The evaluators, though that's not the label they use, think it's obvious that this is a—or perhaps *the*—key ratio needed to appraise the programs, and one that the FDA should require them to release. On this issue (possibly for the first time in my life), I find myself taking sides with the vendor against the would-be consumer advocate, and I think the issue has extremely general applicability. My take is that the key issue is whether the program, *if followed,* will produce the claimed results; and that following the program is (largely but not entirely) a matter of strength of will. Failure to stay with the program—that is, attrition—is therefore (largely but not entirely) a failure on the part of the subject not the program, and the program should not be 'charged' with it.

First, here's why I think this is a very general problem that we need to deal with, in evaluation overall, not only in program evaluation. Think about the evaluation of: any chemical drug abuse program; twelve step programs like AA for alcohol and gambling abuse; distance or online education; continuing education of any kind—this clearly applies to all of them. Now it also applies in some important cases outside program evaluation, ones that you might not think of immediately. Here are two: (i) it applies to standard pharmaceutical drug evaluation because there is a serious problem referred to as the fidelity or adherence problem, about the extent

to which patients ex-hospital do in fact take the prescribed dosage on a regular basis. In these studies we surely want to say that the merit of the drug lies in what it does if it's used, not whether it's used. Case (ii): in teacher evaluation, although we want to say that the teacher has some obligation to inspire interest, to motivate, as well as to teach good content well, success is clearly limited, not only by natural capacity—as we all agree—but also by dogged disinterest. We don't want to blame teachers for failing to teach inherently capable students who are determinedly recalcitrant, i.e., for high failure ('attrition') rates where the cause is simply refusal to try.

Here's the schema I recommend for dealing with this kind of consideration. Think of a program (or drug regimen, or educational effort) as having three aspects that we need to consider in the evaluation: (A) Attractive power; (B) Supportive power; (C) Transformative power. For short: Appeal, Grip, and Impact. A is affected by presentation, marketing and perhaps allocation, and controlled by selection. The vendor or provider has the responsibility to use selection to weed out cases who are demonstrably unsuitable for the treatment; but, given the unreliability of such selection tests in the personnel area (pharmacogenomics is the subject devoted to this in the pharmaceutical area, where it's considerably more successful) and the importance of giving people a chance when they want to try, one can't be very critical of high-pass filtration for weight-loss, distance ed, and twelve-step programs. Of course, high front-end loading of payments may be excessive, if there's no money-back guarantee.

B is affected by support level including infrastructure (e.g., equipment, air conditioning, counseling), continuing costs (including opportunity costs and fees), and ease of use, for all of which the program is largely responsible; but of course B is also controlled by strength of will. If the support, costs, and ease of use are

disclosed in advance and are both reasonable and delivered as pictured and promised, willpower becomes the controlling variable. Which leaves C, the Impact issue, the real kick in the program: will it deliver as promised if we do our part, taking the pill, doing the homework, getting to the meetings? That's the key issue. While the good evaluator absolutely must check to see if the provider has indeed provided what was promised, and that what was provided was about as good as can be provided at the cost level in question, the rest is up to the subjects. Under these conditions, easily checked and often met, attrition is your failure, not the vendor's.

This is an important issue because it's important that evaluation not assume that these treatments are done to people, and are at fault if they don't work. The fact is that they are selected by people as something *they* will *undertake*, *not undergo*, and failure is often the fault of the people not the program. Even with drug treatments, the drugs have to be taken, and often taken for the rest of your life. They only work if you make them work. This is not surgery, which you *do* undergo, which *is* done to you; it's something where you choose to get some help in doing something to yourself. You have to take responsibility for doing your part, and the evaluator must not take that responsibility away and say that the program failed if it didn't get you through to the Promised Land, when it was you who failed. We have free will, but that doesn't mean *success* is a free lunch. Free will is the freedom to start a program: will power is what it takes to complete it.