The Epistemology of Randomized, Controlled Trials and Application in Psychiatry

Derek Bolton

Abstract: The epistemological principles underlying randomized, controlled trials and evidence-based medicine generally have not received the attention they require. Broadly speaking, they are the application of work done over several centuries in philosophy and scientific method. The epistemological base is sound, but it also implies internal limitations, having to do with decreasing generality, which particularly affect application to psychological problems. The principles also have nothing to say about values. The question of the 'objective validity' of scientific method is briefly discussed.

Keywords: Evidence-based medicine, epistemology

Randomized, controlled trials (RCTs) are generally taken in the context of evidence-based practice to be the gold standard of evidence of treatment efficacy. More accurately, accumulated evidence from several or many RCTs meeting stringent criteria subject to systematic review is the platinum standard (if we use what I think is the grading of metals used by some of the charge card companies).

In this paper, I review the epistemology of RCTs, by which I mean the principles by which they are supposed to reveal evidence of causal connections. This is a matter of philosophy of science, the philosophy of causality specifically, and it is has been somewhat neglected in the scientific, clinical, and social theoretic critical literature related to RCTs and evidence-based practice generally. Having sketched what I take to be the strength of the epistemological principles underlying RCTs, I draw out some implications for their limitations, some of which particularly affect psychiatry. Finally, I consider the status of 'epistemological principles' in the current social constructionist climate.

Claims about the limitations of RCTs are widespread and familiar. They can be grouped under various headings, such as the following:

- Lack of ecological validity: atypical participants, atypical clinicians/therapists. This is a familiar charge in the world of psychological therapy RCTs.
- Individuals, contexts, and interventions are in practice (usually) so complex that evidence from RCTs hardly helps, for example, for consumers in multiple kinds of adverse situations at the same time; or in stepped care approaches to management and treatment of some conditions.
- RCTs and evidence-based medicine (EBM) may (or may not) be suitable for pharmacotherapy, but psychotherapy cannot be evaluated in this way, because

of the uniqueness of the individuals and processes in psychotherapy, and because of the difficulty in measuring processes and outcomes that really matter (e.g., change in unconscious defense mechanisms).
- RCTs and evidence-based practice generally are service focused, not consumer focused; they do not attend primarily or at all to the kind of outcomes that matter most to consumers.
- Social power influences, benign and otherwise, the latter including—at the crudest—various kinds of suppression of unwanted results.

Faced with all these formidable objections and limitations, it is a wonder that RCTs and EBM generally don't just pack their bags and go home, back to the pharmaceutical companies and university ivory towers from whence they came, but then, they do have one—just one—claim to major strength: They are meant to provide the evidence for what works—including for conditions that are life threatening or life near destroying—and this in the context, for all but the very wealthy, of limited financial resources for health care. So there are big stakes here, and this is obviously why these matters derive and get much attention.

But is this one major claim valid? Is it the case that RCTs provide the best evidence we can muster for treatment efficacy? As indicated, I interpret this question as an epistemological one, a matter for the philosophy of science and of causal explanation specifically. This is an unusual approach these days. The idea that epistemology and philosophy generally can find an objective, rational, even a priori basis for knowledge is widely regarded as no longer credible, and for good reasons, which all have to do in one way or another with the relativity of knowledge claims. What remains of epistemology is the study of relativized systems of knowledge claims, with their own criteria of evidence, and the power structures that maintain and promote them, a study that belongs not to philosophy, which has dropped off this particular map, but to the social sciences. In this spirit, sociological critiques of RCTs and EBM generally can highlight the power relations at work in the definition, construction, and use of the evidence base, involving professional and commercial interest groups. In this paper I do not want to deny these points, but wish just to argue the case that, nevertheless, there is a definable scientific method for tracking causal connections that has an objective validity.

Hume's Analysis of Causal Connections and Subsequent Elaborations

The best place to start is with Hume's analysis of causality based on generality in the *Enquiry* (1777/1902). In brief, '*A* causes *B*' implies that events of type *A* are always followed by events of type *B*. Knowledge of causal principles thus enables prediction. It also, crucially, enables intervention, because if the production of *A* is within our power, then so is the outcome *B*.

Some decades later Mill, in his *System of Logic* (1843), saw that in practice what is observed on any one occasion is not simply an event of type *A* being followed by an event of type *B*, but this conjunction in a complex of circumstances, C_i. To establish a causal link between *A* and *B*, the possible confounding effects of C_i have to be determined. This involves observing the effects of C_i without *A*, on the one hand, and *A* without C_i on the other. Mill called these principles the 'methods of agreement and difference,' and they underlie our modern idea of *controlled experimentation*.

Further complexities were introduced at the beginning of the twentieth century, in the context of the biopsychological sciences, first agriculture and later medicine, psychology, and social sciences. In these domains, we rarely find universal generalizations, but rather partial ones of the form: A *is followed by* B *in a certain proportion of observed cases*. One function of the Humean universal generalization is to license the simple inductive inference: *The next observed* A *will be followed by* B. In the absence of a universal generalization, the problem is to determine the *probability* of the next *A* being followed by *B*, given that the proportion in the sample so far observed. This is the problem for the *theory of statistical inference*. Its complexity compounds in interaction with the problem cited, that of controlling for potentially relevant confounding variables.

Relevance to EBM and RCTS

So far as concerns evidence in health care, these methodological principles point toward–with some intervening steps—the *randomized control trial*. The trial seeks to determine as precisely as possible the 'active' causal component of the intervention and the condition that is affected. This means that other associated factors must be controlled for, both in terms of control groups and control interventions. The scientific paradigm can countenance study of individual cases, but the pressure is always toward generality and hence groups, these moreover of size sufficient to detect the effects of interest according to statistical power calculations. Statistical methodology has had to be adapted to cope with the particular demands of evaluating treatments for human beings. Much of the methodology was originally worked out in relation to relatively controllable events, such as in agriculture. In clinical trials, on the other hand, we have to countenance potential participants deciding they do not want one arm of the choice of conditions, or participants walking out half way through the trial. Such eventualities demand further statistical sophistication, understood and developed by statisticians and other trial methodology *cogniscenti*.

A technical aside is necessary here to comment on the 'intervening steps' that lead from the methodological principles due to Hume, Mill, and the theory of statistical inference, to the RCT. The steps have to do with the methodological assumption that randomization serves to control for potentially confounding variables by ensuring that the characteristics of the groups receiving, for example, active treatment and placebo are the same (with respect to the potential confounds). Obviously, randomization does not guarantee this identity, although the probability that it does achieve it is increased with the size of the groups. If major potential confounds are identified (such as, depending obviously on the condition and treatment, gender, age, or comorbidity) then stratified randomization may be used to ensure that rough equal numbers individuals with the potentially confounding characteristic are randomized to the two trial conditions. For example, participants recruited into the trial are divided first by gender and then randomized to active treatment or placebo within gender groups, ensuring that (roughly) equal numbers of men and women end up in each arm of the trial. The principles and procedures here are dealt with in the trial statistical literature (e.g., Pocock 1983/1995).

These considerations also point to the fact that, although RCTs are the gold standard for treatment efficacy, there is a gradation of strength of evidence, beginning with quasi-experimental single-case studies, moving to case series, to 'open,' uncontrolled trials, through to RCTs. With each step, there is increased generality and increasing control of potential confounds. This approach of defining hierarchies of evidence base, culminating in systematic reviews or meta-analyses of several or many (well-designed) RCTs, is of course the one adopted in treatment guidelines such as issued by NICE in the United Kingdom. It may be noted also that what typically is described as the lowest kind of evidence base, expert consensus, is not continuous with the quasi-experimental and experimental designs referred to above; it is a different kind of evidence altogether, invoking expert custom and practice, nothing to do with the scientific investigation of causes.

According to this account, then, RCTs have an epistemological basis, the development of which stretches back over at least some two hundred years, resting on Hume's insights about causality and subsequent sophistications and adaptations. These are the principles. The question whether they work is basically a pragmatic one: If this methodology comes up with the result that treatment *T* cures condition *C*—more likely, that *T* raises the probability of remission of *C* from a base value to a higher one—is that true? This can be answered by subsequent applications of the treatment, including in other RCTs. All this is a blend of British empiricism and American pragmatism, and this gives a hint as to where the major philosophical alternatives are to be found.

Complications in the Social Sciences and the Particular Position of Psychiatry

There is in fact a well-known and philosophically profound problem with applying the Humean

notion of causality to human beings and along with that scientific method generally. I refer here to the meaning/causality schism that appeared some hundred years ago in Continental Europe and remains active.

The end of the nineteenth century saw the appearance of a new kind of science, called in German the *Geisteswissenschaften*, including history and social science, that have as their subject matter the expression of mind in society, and the *meaning* that permeates human activity. With the appearance of these sciences, there arose a fundamentally new problem, which remains ours, namely, that knowledge of mind and its expression in meaningful activity does not conform readily to the methodological assumptions and rules of the natural sciences, and requires instead a distinctive hermeneutic methodology. The tensions that found expression in the celebrated distinctions between *meaning* and *causality*, and between *understanding* and *explaining* included the following:

- Whereas natural science deals (mostly) with repeated and/or repeatable phenomena, an historical event—such as the decline and fall of the Roman Empire—or a cultural practice such as a particular system of government—are singular, or even unique.
- Whereas natural science seeks and uses general causal laws in its explanations, history and social science construct and use diverse ways of understanding a whole variety of events.
- Understanding seems to be subjective, to draw on empathic abilities, which vary from person to person, or from culture to culture, whereas the methods of observation in the natural sciences are objective, and the results are meant to be the same for all.

The whole problematic hit psychiatry harder than most disciplines, mainly because it never had a choice but to embrace *both* hard, causal science, *and* the need to understand so far as is possible abnormal states of mind and the person who has them. Jaspers (1923/1963) was the first to grasp the relevance of the new problematic to psychiatry, and perhaps the last to be able to hold on, even handedly, to both methodologies, emphasizing the importance of both the science of psychopathology and the indispensable need for understand meaning by empathy.

Another pioneer in psychiatry, less of a phenomenologist and more of a scientist than Jaspers, came across the same problem in another way, and indeed pointed way ahead to a solution. Freud saw that some apparently senseless mental states and behavior could be understood as meaningful, and that intervention in the meaningful processes could effect change, that is, would be causal (e.g., Breuer and Freud 1895/1978). Freud the neurologist recognized that if this were so, then the mental, meaningful processes would somehow have to be mapped onto, realized in, brain processes. But how this would be so, what the architecture and functional characteristics of the brain would have to be like for it to be so, were questions that Freud recognized could be answered in the then present state of cognitive neuroscience, and he left his 'Project' unfinished (Freud 1895/1978; Kitcher 1999; Spitzer 1998).

I have argued elsewhere (e.g., Bolton 2003; see also Bolton and Hill 2004) that something like Freud's project is on-going in the current information-processing paradigm in the behavioral sciences, including current models of psychopathology, but the focus in this paper is on causality and the basis of RCTs. In this context, the main question is whether there are generalizations about human beings, which would enable application of Humean principles of causality, made considerably more sophisticated as above, or whether individuals are unique, for whom hermeneutic methodology is fitting.

The answer to this question whether individual human beings are unique is, I suggest, a truly postmodern one, namely, yes and no, depending on a variety of considerations and contexts. *No* (not much), in relation to, for example, orthopedic surgery and cancers; *yes* in relation to the total (or large scale) life narrative; *no* in relation to many significant meaningful themes in life narratives, such as *melancholia is normally linked to mourning*; *expectation of noncontingency induces helplessness*; and *people want autonomy over their lives and for others to respect this*. There will always be, I suggest, a dialectical play between uniqueness and generality, between what we do and what we do not have in common, and consequently the blanket, ideological positions—everything human can fall under science of causality, and nothing human (and meaningful) can—are both wrong. I am not sure that this is primarily a matter of meaningful-

ness as such (still less of consciousness), or whether it fundamentally one of systemic complexity and individual differentiation. Meteorologists may have similar issues about weather systems.

To put the point another way, matters human can be graded on dimensions of generality to individuality. To the extent that there is generality, there is the possibility of determining causes using the scientific methodology as outlined (RCTs in medicine and health care generally). To the extent that there is individuality, the methodology loses traction, and with it the concept of causality itself, which transforms into the concept of self-cause, or agency.

In this way, the problems inherent in the determination of causes and RCT methodology in particular hit particularly hard in psychiatry, and especially in psychotherapy. The reason is simply that mental health problems, by and large, in contrast with physical health problems, by and large, do involve the person as a whole, not some subpersonal biological organ, perhaps the total, or at least a large scale life narrative, which tend toward the unique and therefore nongeneralizable. Psychotherapists in the psychoanalytic tradition tend to highlight this feature, which is one reason why the evidence base in this tradition has tended to stay at the level of the single case history. But, on the other hand, as noted, there is in fact no shortage of generalizations in psychopathology and related domains, even, and of course, from the standpoint of psychoanalytic theory itself, such as the links between melancholia and mourning, or, to update the examples, between severely disrupted early attachments and 'attachment disorder,' and so on. Generalities of this kind about causal psychological mechanisms are viable hypotheses, as are related hypotheses about which interventions will or will not be effective in prevention or amelioration, and they can all be tested, including in intervention studies. On the other hand, again, to the extent that particular conditions or experiences really are unique to an individual, then predictions made on the basis of previous experience—and probably understanding itself—will lose grip.

The determination of causes and RCT methodology in particular are limited by internal considerations, by their own methodological principles. The methodology depends essentially on the occurrence of repeated and preferably repeatable events to determine which among the various factors involved are causal, and to what extent, and in order to make predictions. Scientific methodology is therefore not a good fit with relatively rare events involving many potentially causal factors. An example familiar in health care provision that is jeopardized in this way is the question of relative effectiveness of different models of health care delivery, such as in-patient and community care, or of two kinds of community care. In this kind of case, the phenomena being compared are so complex, involving so many potentially causal factors, while at the same time, and connected, the capacity to observe repeated phenomena is so limited, that scientific evaluation, and RCTs in particular, are likely to be problematic. This is not to say, on the other hand, that scientific method has no role in comparisons of this kind, but it depends on breaking up very large questions into smaller, more manageable ones, where the requirements of repeatability and control over the number of potentially relevant variables can more easily be met. In evaluating a whole system of community care, for example, it might be possible to conduct trials evaluating particular components, such as team care as opposed to care continuously by a particular member of staff.

The point that traditional scientific methodology is not a good fit with relatively rare events involving many potentially causal factors was of course precisely the insight of early theorists in the *Geisteswissenschaften*. Social systems do not readily lend themselves to repeated observations, and they certainly involve many factors. Decisions here tend to be made on grounds other than evidence; evidence-based policy tends to give way to policy-based evidence: Decide what you want to do, or what you want to believe, and then go looking for the evidence (and ignore the counterevidence). The question of the scientific study of human beings, as argued, is always ambiguous in this respect. It is possible to argue on the one hand that each human being is unique, and certainly complex, and on the other that we have much in common, and there are certainly many of us, so that repeated observation is quite possible. One side of this dialectic points

to hermeneutics as the appropriate methodology for understanding human beings; the other leads to the psychological and behavioral sciences, one expression of which is the principle of using RCTs to guide treatment choice in health care.

One limitation of RCT methodology, then, is the internal problem of generality, clearly signaled early on in the meaning/causality distinction. The other limitation of this scientific methodology is external, and nothing obviously to do with that distinction. It has to do with questions of values.

The Epistemology Underpinning EBM and RCTS Has Nothing to Say About Values

Scientific method—insofar as it keeps fidelity to the method—is value free. This is so notwithstanding the fact that it is used by individuals and structures that are pervaded by values and concerns. The scientific method for tracking whether *A* causes *B*, whether a particular treatment effects a particular condition, or has particular side effects, it entirely neutral to a whole range of questions of value, such as:

- Is this an interesting question?
- Is the research to answer it worth funding?
- Is the condition of interest?
- Do people want the treatment?
- Are the effects of interest?
- Is it better to devote limited financial or staffing resources to the relief of current extreme suffering of the few, or to a public health preventative program for the many?

Science can contribute to the assessment of some factual questions, such as the assessment of consumer satisfaction, but is neutral to the moral questions themselves. And these questions are not tacked onto to the science when the science is done; the science cannot even start until they have been answered.

Concluding Remarks on Scientific Method and Social Construction

Finally, let me return to the methodological problems mentioned at the start, the question of the prospect for an objectively valid scientific method in the current social constructionist *episteme*. Social science seeks to understand social change, and whatever else it may be, EBM is certainly that. From the sociological perspective, power relations and relativity are open to view in all social change. From this point of view, we see the interests and maneuvering of the pharmaceutical companies that shape, interpret, and publicize the results of treatment evaluations for the benefit of, among other things, their sales; we see the power that accrues from the credibility base of EBM, the way that it influences allocation of clinical and training resources; and all this in the context of recognizing that one group's 'evidence' is another's irrelevance, that each epistemological paradigm defines its own, the scientific preoccupied with the objective and generalizable; the 'holistic' or 'hermeneutic' paradigms with the subjective, the intuitive, the unique. EBM appears in this way as the victory of one knowledge paradigm over another, helped along by the vested interests involved, but no more objectively valid. In the constructionist social science paradigm there is little or no time for the notion of objectively valid, rational principles of evidence.

Philosophy, on the other hand, does, or least did, believe in such a thing. The demise of the credibility of this idea corresponds roughly with the increasing dominance of the constructionist social science paradigm, from roughly the 1970s. The demise of the idea of objectively valid rational principles of evidence has been most immediately the collapse of the idea of scientific method. One history in the philosophy of science includes the following milestones. In the 1920s, the logical positivists gave a clear late expression to the principle of induction as being the defining characteristic of empirical reasoning and empirical science. Right or wrong, Popper (1959) dethroned induction, emphasizing instead the falsifiability of scientific hypotheses, although maintaining at least the idea that scientific theories were answerable to objective empirical evidence. However, subsequently developments in the Popperian school wore away at these alleged characteristics of scientific method. Notably, Kuhn (1962) introduced the idea of 'paradigms' that, crucially, defined their own observation statements (evidence). From

there, it was an easy fall off the log to the relativity of epistemology, to the end of epistemology as practiced in the tradition, leading to titles such as Feyerabend's *Against Method* (1975), which emphasized social power relations in the history of science—a philosophical expression of the social constructionist paradigm that was flowering at the time.

This death of epistemology (in the traditional sense) happened, ironically, at around the time that EBM was taking off, in medicine generally and subsequently in psychiatry, the first trials of prescribed psychotropics and of behavior therapy in the 1970s. Irony indeed that medicine and psychiatry began to make such use of scientific methodology right at the heart of clinical practice just when the cultural shift was toward relativizing epistemology. And by now EBM has reached a stage of maturity, with the accumulation of many treatment evaluations, so that we have professional body or government agency stamped clinical guidelines, advising, dictating, or in any case prioritizing what treatments should be given and which will be paid for. But is this just the victory of one epistemological paradigm over another, of one set of power relations over another, riding roughshod over other forms of clinical knowledge? Or is it an objectively valid advance?

One of the advantages of considering scientific method in relation to EBM is that here the method is seen clearly as having practical relevance. As in all 'technological' contexts, science appears less as a claim to knowledge of a universally valid truth, and more as a tool to achieve certain results. It is possible to construe Hume's 'analysis of causality' as being a formulation of what kind of evidence and reasoning is required to facilitate the achievement of desired goals by action. As such this epistemology is of interest to all agents and in this sense it is 'universally valid'; it appears as sophistication of folk empirical reasoning. To put the point another way, all agents are interested in knowing what differences their actions will (are likely to) make, and what they should do to achieve particular desired outcomes. It is plausible to say that this interest is found also in familiar epistemologies that are usually contrasted with the scientific, such as in magic, where rituals are intended to have certain practical effects, and in some religious practices, such as some uses of prayer, for example; and in the 'nonscientific' psychotherapies, which after, all, in some sense, seek to be able to foster change (of some kind). So it may be the case after all that scientific method has a universal rational status, even though, and indeed definitely in the midst of, all the other factors that guide and affect individual and social action.

References

Bolton, D. 2003. Meaning and causal explanations in the behavioural sciences. In *Nature and narrative. International series in philosophy and psychiatry, volume I,* ed. K. W. M. Fulford, K. Morris, J. Z. Sadler, and G. Stanghellini, 113–25. Oxford: Oxford University Press.

Bolton, D., and J. Hill. 2004. *Mind, meaning, and mental disorder: The nature of causal explanation in psychology and psychiatry*, 2nd ed. Oxford: Oxford University Press.

Breuer, J., and S. Freud. 1895/1978. *Studies in hysteria*. In *The standard edition of the complete psychological works of Sigmund Freud*, vol. 2. London: Hogarth Press.

Feyerabend, P. 1975. *Against method*. London: New Left Books.

Freud, S. 1895/1978. *Project for a scientific psychology*. In *The standard edition of the complete psychological works of Sigmund Freud*, vol. 1, 283–397. London, Hogarth Press.

Hume, D. 1777/1902. *An enquiry concerning human understanding*, ed. L. A. Selby-Bigge. Oxford: Oxford University Press.

Jaspers, K. 1923/1963. *Allgemeine Psychopathologie*, trans. J. Hoenig and M. W. Hamilton. Berlin: Springer Verlag. English translation by J. Hoenig and Marian W. Hamilton. *General psychopathology.* Manchester: Manchester University Press; 1963.

Kitcher, P. 1999. Sigmund Freud. In *The MIT encyclopedia of the cognitive sciences*, ed. R. A. Wilson and F. C. Keil, 328–9. Cambridge: MIT Press.

Kuhn, T. S. 1962. *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Mill, J. S. 1843. *A system of logic*. London: John W. Parker.

Pocock, S. J. 1983/1995. *Clinical trials*. London: Wiley.

Popper, K. 1959. *The logic of scientific discovery*. London: Hutchinson.

Spitzer, M. 1998. The history of neural network research in psychopathology. In *Neural networks and psychopathology*, ed. D. J. Stein and J. Ludik, 88–131. Cambridge: Cambridge University Press.