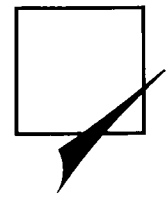
Evaluation

[1356–3890 (199707) 3:3]
Vol 3(3): 363–374



The Debates, Notes and Queries section provides an opportunity for evaluation-related issues to be debated as well as for more general interchange. Debates can take the form of sustained arguments by the advocates of different approaches or of briefer thoughts or notes. Contributors may also wish to comment and raise questions about material that has previously appeared in the journal or simply use the section to draw readers' attention to relevant issues, ongoing research, evaluation activities and other events.

Consumers and Heroes: A Critical Review of Some Recent Writings of Michael Scriven

As one of the patriarchs of evaluation Scriven clearly has an interest in its fundamentals. Some of his recent writing (Scriven, 1993, 1996a, 1996b) has been reviewing what happens to some of those fundamentals as evaluation matures into an independent discipline. The last two of his addresses were directed to, respectively, the British and American Evaluation Associations through their journals. The other is a monograph in the prestigious series New Directions for Programme Evaluation—noteworthy, for each issue of this series usually contains articles from four or five authors, whereas this one is devoted entirely to Scriven. All of these are significant, for these journals (both of high quality) and the monograph are primary sources of reference for new and young evaluators—and it is upon these neophytes alone that we depend for better versions of evaluation theory and practice than we have so far managed to discover. Now when a tribal elder speaks we have to pay attention carefully—for the sake of methodological advancement, the balance between respect for inter-generational continuity and an essential intellectual autonomy of the new generation is a precarious one indeed. (No wonder Zen teaching advises us to 'kill the teacher'.) So Michael Scriven must be listened to—but sceptically.

My main concern with respect to these articles—the basis for my scepticism—is that Michael Scriven is falling behind the times and that this is revealed in these showpiece articles. This might be thought to be an unusual charge to level against an evaluator of Scriven's experience—after all, what is an evaluator supposed to be doing all day long but immersing him or herself in current events in real time? But this is to discount the inertia of success and biography—Scriven has a lot of both; and it discounts the vulnerability of all evaluators to cooption by the logic of the day which suffuses social programmes and which inures the evaluator to contemporary social

change. Even so, what marks out evaluation as a form of enquiry is the necessity of its engagement with social change. This presents a problem for older generations of evaluators (I count myself in an intermediate generation) who suffer a corrosive fear of change—of 'altered' or 'unstable' states, as Schon (1971) argued long ago. Innovation observers have their own problems with change.

Evaluation practice, new a discipline as it may be, is changing in concert with the ebbing away of liberalism in our democracy and the collapse of traditional forms of social organization (Adelman, 1996). Scriven's concerns with robust theoretical underpinning for evaluation practice (Scriven, 1996a) and with the demanding set of competencies that define the eponymous 'evaluator' (Scriven, 1996b) sit uneasily with the reality that much evaluation is currently commissioned from anti-intellectual firms of management consultants who trade in formulaic solutions; that they and sponsors often hire solitary and vulnerable individuals to conduct the investigation; that much of evaluation methodology is held to be 'commercially sensitive' and, therefore, secret; and that the intellectual 'upstarts' of constructivism and postmodernism have scored some undoubted successes in engendering mistrust of conventional approaches to theory and role. His concern with the overriding function of evaluation to inform the 'consumer' (Scriven, 1993) ignores the considerable debate, somewhat sparked by the recent successes of neo-liberalism and market theory, about the lack of fit between the twin roles of consumer and citizen, and ignores the extensive movement across the professions and public services to address issues in research-based and community-based practice.

In this essay I want to concentrate on by far the most important of these pieces of work, his 1993 monograph, and as I do I will focus increasingly on the relationship he advocates between evaluator and what he calls the 'consumer' of public programmes. Most references which follow are to the monograph.

Of Adolescents and Elders

The field of evaluation, one might infer from Scriven, is roughly in its adolescence. One might, therefore, expect from it a degree of rebellion, prodigality—youthful contempt, even. Nor would you be surprised to hear disciplinary talk from the older generation. So it is that Scriven sets himself the task of reviewing the wanderings of this uppity child and—with really little hint of regret—whipping it back into line. Here is the smack of authority, impatient with stragglers and dissenters, written by an undisputed 'elder' of the evaluation tribe with a licence for irascibility. He reasserts the importance of evaluation and its demands for sophistication in practice. The 1993 monograph is, in fact, useful reading for understanding foundational values in evaluation and how they are often abused, offers clearer expressions of key concepts and practices than are generally available—even, as I shall explain, provides a good example of how methods of enquiry reflect the zeitgeist (House, 1993). Scriven talks in practical terms of the status of evaluation as an autonomous field and profession and effectively dismisses the current disposition of the social sciences to do the work of the evaluator—and, therefore, the erosive influence of social science concerns on evaluation practice. He reminds us that 'establishing statistical significance is the easy part of establishing significance' (1993: 70). Here is good retraining material for those

economists who think that programme evaluation is merely the application of value-free technology.

Throughout these articles Scriven offers an analysis of evaluation as a 'transdiscipline'. Here, evaluation takes its place as a generic field of instrumentation for enquiry on a par with statistics and logic. Evaluation can 'provide major tools for at least some disciplines' (1996a: 403) and, in common with logic, is one of the only two processes of enquiry that all disciplines share—Scriven casts the 'transdiscipline' of evaluation as a process of reflexivity which all disciplines engage in to monitor their disciplinary boundaries.

Fair enough, though this offers a rather conventional scholastic view. The combination of philosophy, feminism and discourse analysis, for example—sometimes in the form of postmodernism—has recently wreaked an iconoclastic vengeance on disciplines as diverse as architecture, political theory and education. That was the right hook. The left hook came similarly from constructivism, the errant protégé of the philosophy of science. Perhaps these might be thought of as forms of evaluation, in Scriven's terms, offering resources for boundary-reflection, but that would stretch the definition of evaluation far beyond what Scriven would countenance. The point is that the concept of a transdiscipline is not uninteresting, but applied to evaluation it probably owes more to an evaluator's self-esteem than to correspondence with any significant reality. I certainly see little evidence of people resorting to evaluative procedures to help accommodate themselves to change or to reflect on the nature of their own disciplines. Flight, conflict, subversion and retrenchment are more common strategies among those I evaluate in contexts of university change, for example (e.g. Kushner & Marfleet, 1995 is an evaluation of a university innovation in modularity and credit accumulation). In any event, Scriven sees status as a transdiscipline as providing opportunities for evaluation to extend its methodological range—but there is little evidence of this in his own prescriptions which are conventional.

Joining the contemporary backlash against relativism and constructivism, Scriven seeks to reassert a degree of objectivism—it is, he argues, perfectly possible to assess the merit or worth of a programme and measures of these things are 'real'. Evaluation for Scriven is a process of determining what he calls in the 1993 monograph 'certain evaluable properties of things'. I, too, believe that social constructionism has to have limits in the field of policy and programme evaluation (Kushner, 1996) because I also believe that there are what we are obliged to take as objective social realities—ask the poor, the sick and the elderly after 20-odd years of monetarist experimentation. But Scriven's treatment of objectivity (in this monograph) is light compared to that of co-backlasher Phillips (1990), for example, who argues that objectivity is more sustainable (and would win more supporters) where it can be distinguished from certainty and if it is seen more as a set of properties in enquiry than as the end point of a realist quest. Scriven's response to constructivism is simply to ignore it (though in 1996a he dismisses it by referring to the lunatic fringe of extreme individualism—implying Guba and Lincoln, 1989). Real measures of programme worth are merely 'complex' (interestingly, the word 'controversial' does not appear in the monograph) but nonetheless attainable. The difficulties we evaluators face each day in responding to

value-pluralism are barely addressed and not at all eased by Scriven who, in these recent writings, does not acknowledge it as an issue.

As part of the same redemption programme he reasserts in the monograph the evaluator's obligation to synthesize data and analysis—i.e. construct arguments—implicitly taking on those who would argue that this is merely textual appropriation. In this he is with House (1995) who argues that evaluators can reach 'all-things-considered judgements' without lapsing into philosopher-kingship on the one hand or relativism on the other. Scriven and House share a distaste for the lessons evaluators have learned from social science which persuade them into value-free stances. These are not, as Scriven reminds us, in the trivial sense that a researcher must not allow personal preference to intrude, but in the more significant sense that social enquiry deals only with descriptive and analytical matters, not with the normative—i.e. they deal with what *is* and not with what *ought to be*. Scriven is impatient with such a view arguing, among other things, that the scientist who promotes value-free enquiry happily exercises conclusive judgements in other social guises. House goes a little further than Scriven in arguing that evaluators also have the warrant to weight and balance interests in analysing evaluation data—for example, to give greater weighting to disadvantaged groups in a Rawlsian strategy of positive compensation.

In fact, Scriven makes no reference to issues of social justice at all, although he does talk of the need for evaluators to rediscover the role of normative judgement. Ericson (1990) is, too, aligned with the movement back to some foundational principles that have been eroded by value-free social science and which are back on the agenda with the titanic battle being waged between utilitarianism, libertarianism and the Rawlsian 'justice-as-fairness' view. He sees the growing acceptance of the underpinning of evaluation by principles of social justice in relation to the distributive dynamics of society according to one of these three approaches. 'An evaluation that simply attempts to identify in a summative or formative manner actual or likely programme consequences is necessarily incomplete' (1993: 19). Scriven talks of 'Absolute values' (1993: 30) bemoaning their absence from standard texts on programme evaluation. But when he comes to provide examples, he translates absolute values into absolute ethical principles and these into day-to-day moral standards. He urges the evaluator to report on programme managers stealing from the till, sexually harassing office staff and running an illegal book. Barely a threat to the neo-liberalists and other 'terminators' of social programmes.

So, too, in, for example, one of his two articles (1996a) in which he rightly bemoans the failure of 'practical' evaluation approaches to employ theory. But here, too, Scriven is more concerned with protecting the fundamentals and conventions of evaluation than, say, with questioning its epistemological or political warrant. Being theoretical, for Scriven, is being aware of the sources of evaluation instrumentation; coping with challenges to validation based again on the deployment of instruments and processes; researching evaluation roles; being aware of sources of bias; and being aware of the range of evaluation styles. These are concerns *internal* to the practical culture of evaluation and do not reflect any preoccupation with the changing external environment. To illustrate the gap, Bob Stake not long ago came to my Centre and started a talk with the words 'it's over'. Like his father's occupation, responsive programme

evaluation had enjoyed a historical window of opportunity the span of a single generation. The Reagan/Thatcher years—more generally, the inexorable rise of managerialism—had sealed its end. This, indeed, is evaluation theorizing but represents a level of reflection barely touched by Scriven who is quite literally 'out' of time.

Rooted in past foundational values this monograph does not move evaluation far forward—much less seek to learn from some of the recent adventures of the evaluation adolescent. Scriven's purpose is to enquire into 'the mistakes we have made' and to discover consequent learnings, but eventually he achieves this within much the conventional set of values and assumptions that are all too familiar. In the monograph he downgrades process, portrayal and formative evaluation in favour of comparative, summative and product evaluation ('formative evaluation is attractive, but summative evaluation is imperative', p. 59)—although in one of the articles (1996b) he recovers the position by both reminding us that the formative/summative distinction is far more complex than normally assumed and by protesting that the preference for formative and summative is dictated by circumstance and not by evaluator's preference. However, he confuses the issue further by arguing that both summative *and* formative evaluation are about determining the 'merit or worth' of a programme. In the end, then, formative collapses back into summative—'product'—evaluation.

False Surrogates for Programme Evaluation

Scriven's justificatory appeals in the monograph are telling. A frequently invoked solution to dilemmas and conflict is the *needs analysis*, which by implication serves to displace evaluation and programme controversies to a place where they succumb to rational, empirical arbitration. A needs analysis for Scriven is more than asking people what their needs are and serves as a surrogate for what many others would include in the programme evaluation itself. But there is an analytical flaw in this displacement. He is fond of reminding us that people who wage war against values in one role are passive and accepting of them in others—which suggests that they are not consistent in holding to their values across social contexts. Why should they be? But in such a world where our values and beliefs were context-independent (I do not live in that world) value conflicts over, for example, the meaning and esteem of a programme's impact would be consistent as between the programme evaluation and the needs analysis—however sophisticated—which supplemented it. Social programmes tend to evoke controversy over core social values, the kind which would arise elsewhere as a matter of course—for example, that the state should/should not interfere in people's lives; that autonomous activity on the margins of programmes represents creative and desirable diversity/chaos and unfaithfulness to policy. Surely the whole point of evaluation is that such differences cannot simply be arbitrated away by surrogate studies.

Confusing Consumers (not Shareholders) with Citizens

More serious an issue, he constantly and consistently invokes 'the consumer' of public programmes to discipline our sense of methodological licence and our more esoteric preoccupations. His approach, he says, 'regards the consumer's welfare as the primary justification for having a programme and accords that welfare the same primacy in the

evaluation' (1993: 9). This he asserts to refocus evaluation away from its now conventional orientation to supporting programme decision-making. With this I have no argument—with Scriven's notion of how to enhance the welfare of the consumer, I do.

The consumer—the tax-payer, the parent, the patient—he says is not interested in the niceties of process or of programme improvement techniques, s/he is interested in whether the product is better than another and whether it does what it is supposed to do. The consumer is interested in value-for-money (hence Scriven's repeated injunction—with which I also sympathise—to train evaluators in cost analysis). 'Most consumer or taxpayer groups have little interest in whether a programme meets its goals as such, only in whether it does something that needs doing, whether the cost is reasonable and whether it does it better than alternative ways of doing it' (1993: 20). Here is the economic-paternalist view of the citizen—a view which Plato might well have held were he trained in a contemporary low-rent school of economics. Paying your tax pounds or dollars entitles you to little other than the right to buy or to refuse to buy. Certainly, there is no sense in which paying taxes gives the citizen ownership rights over social programmes. Indeed, we may note that the word Scriven chooses is consumer, which implies a contractual relationship based on mutual obligation and duty, where he might have chosen the word 'shareholder' which implies accountability and rights—and is, surely, a more accurate characterization of a taxpayer in a democratic state—if you insist on using the language of business interactions to talk of somewhat more complex phenomena. But the use of the term consumer makes Scriven thoroughly modern. Here is an approach to evaluation designed nicely (as the engineers say) for the market.

I want to question this orientation and to do so I will choose some examples that Scriven uses to question his interest in programme products and summations. He says, for example (1993: 59), 'while life-and-death matters can hinge on good summative evaluation, however, we would only have fewer good options to choose among if there were no formative evaluation', and he mentions drugs. In the UK it has long been recognized and published by medical researchers that the simple procedure of doctors prescribing aspirins to those vulnerable to heart and circulation complaints would have a significant effect on sickness and mortality rates. Dissemination and education programmes have barely produced any response from practitioners. Why? Incredibly, we simply do not know. We may speculate, of course, that it is to do with professional habits and psychological resistance, with the popular image of the aspirin potentially undermining the doctor's status—but all we can be sure of is that the only route to effective implementation of this highly beneficial research finding is to understand the individual doctor—many of them, in fact. No doubt there will be no simple answer and a battery of solutions to the problem will be necessary, grounded in a mixture of understanding of individual psychology, professional training, variation in age and gender, variation resulting from the relation between individual doctors and different types of group practices, relations between doctors and drugs companies, variation by social class, etc. There are grounds for believing that all these may be relevant according to context.

Now understanding the basis of professional resistance to alternative prescribing is

well within the bounds of evaluation practice. 'To understand or explain the social phenomena involved in a programme' (1993: 40), though roundly dismissed by Scriven, is the bread and butter of many contemporary British evaluators who still assert their independence of social science as vigorously as he. Indeed, in the UK, educational evaluators are called upon more and more to assist in understanding professional development programmes for the very reason that we are familiar with problems and issues in change, for change is the basic condition of learning. As understanding grows that (post-Weberian) organizational and professional development have to be conjoined, so Schools of Education find themselves increasingly in demand for their educational expertise. Quoting Cronbach and his associates (1980) who say in their major revisionist work that 'Evaluation is better used to understand events and processes for the sake of guiding future activities', Scriven says that 'that is a very appropriate task for applied social science, but it would be suicidal for evaluation to shoulder it' (Scriven, 1993: 40). Much less, I think. It would be economically suicidal—as well as a dereliction of social obligation—for evaluation groups like my own to ignore such tasks.

The point about this is threefold: first, that lives are certainly being lost owing to the lack of descriptive, portrayal-based, process—even biographic—evaluation studies; second, that we have had to wait for the summative results of dissemination programmes before we can even raise these questions. Summative approaches have denied us early solutions.

The third element is to do with Scriven's view of a social programme and derives from the bland assertion that 'after all, programmes are simply ways of getting things done, just as artefacts are devices for getting things done'. If that were so we would 'simply' be getting on with the job of persuading doctors to prescribe aspirin. If we have learned anything from the past 30 years of programme evaluation it is (a) that there is nothing 'simple' about getting things done as is the case with aspirin—this is a breathtaking denial of pluralism; and (b) that social programmes are often to do with anything but getting things done. Social programmes, in the UK at least, are often more to do with inter-departmental rivalry, the personal careers of Ministers of the Crown and their civil servants, the promotion of a vision of economic society, the last-ditch attempt by a community to rescue themselves from economic and social oblivion, expressions of helplessness in the face of technological change and, of course, the management of decline. We have even found teachers joining a 3-year innovatory programme in flight from rapid and continuous change—the programme was an opportunity to 'stand still for 3 years'.

Insouciant Consumer or a Scriven Coopted?

There is, however, a more dangerous view lurking in Scriven's characterization of the consumer in the context of programme (i.e. product) evaluation. His claim is that the 'consumer' has no interest in programme processes—how goals were met—only in the end-products of meeting those goals. The consumer is the very same mindless, rational cipher found in contemporary market theories. How close to the truth can that be? Is it likely to be true, for example, that parents are interested only in their children's assessment results and not in how they are taught in school? Is it the case that we all

turn a blind eye to the vicious means–end rationality that delivers up the tasty two-minute burger on the back of the cynical exploitation of youth labour? Are citizens so obsessed with summations that they are really persuaded by government that the sheer size of the prison population guarantees their protection from crime rather than prison regimes of rehabilitation—or, equally, that it doesn't matter how the police treat incidents so long as the arrest rates are rising? Are we impartial as to the clinical procedures doctors use just so long as they cure us? Are we really unmoved to speculate again on the justness of wars when we watch documentaries that reveal how many soldiers die in training? Are we so numbed that we merely complain when we receive poor or insulting service and suppress our curiosity about the kind of training that provided it?

Who is this consumer anyway? As Scriven says accusingly of scientists, the veil of value-free pretence slips all too easily when the scientists assume their daily life as citizens and consumers ... then watch them change their tune! But is this not so for us all? Evaluators, programme personnel, programme clients have a life beyond 'the programme' and in that life we all transfer learnings, explore social life, become more questioning and sophisticated with age. Knowing what we do at work, do we really shut off when we 'become' the consumer and limit our judgements to the simplest and most self-obsessed? If all these were true we would find little of a social base on which to mount a democracy—but why would we want to bother to mount one anyway? Is not the ultimate vision of the consumer the voting zombie who is supposed to rest content with this year's tax reduction and watch mutely as the cavorting Bourbons cast their wild shadows on the inner walls of secret parliaments? Scriven writ large on the cave walls of our democracy would have the voter insouciant, observing cash-for-questions, politicization of the judiciary, gerrymandering and the like neither seeking nor having the right to make comment.

Is it not more reasonable to think that with all his experience and so many years of focusing on the nature of programme evaluation, Michael Scriven has become a creature of the social and economic programme—so much so that he is coopted into the same restricted language of the programme manager whose very life depends on narrow classifications of groups and their interests and, these days, on market theories of social transaction? In spite of his wanting to shift the evaluation emphasis from programme decisions to consumers of programmes, Scriven propagates the rational, decision-theory view of a social programme—partly, in a nice paradox, by the very invocation of the consumer who has little ontological status beyond the narrow sphere of business decision-making. Consider an alternative. What if we were to measure Scriven's programmes against the meanings and values of an individual consumer/citizen, rather than measure people against the characteristics of programmes as he does? Here, again, is a possible new direction for programme evaluation methodology, receiving something of an impulse from contemporary interest in, for example, biographic research. The argument in this monograph offers Scriven little opportunity to engage in this sort of theoretical extension of programme evaluation, since it lacks the perspective on social justice.

But there is another side to this which is wholly ignored by Scriven. What of the programme professionals—apart from the manager-cipher—the doctor and the police

officer and the service manager? Are they also unconcerned about programme means and programme sociologies? If so, why do we find such ready professional audiences and sponsors of hard cash for the kind of process, formative and portrayal evaluations that seem to address the problems they deal with from day to day? The principal answer is that they seek these approaches—again I say, at least in the UK—because they promise greater cognitive control over social programmes. Another key lesson from programme evaluation over the years and a consistent finding in my own evaluation group is the remarkably low level of knowledge and the concomitant and even lower degree of control exerted by managers or anyone else, for that matter, over a social programme. These people come to us like drivers searching for the steering wheel.

The demands of contemporary professional change are not well served by the kinds of information currently generated by programme evaluators. For example, in the fields of policing and medicine there is growing acceptance of the need to discover both policies and practices—and their legitimation—through dialogue with communities, facilitated by university-based evaluators. There is even pressure in many forms of professional practice which underpin the programmes Michael Scriven evaluates to make those practices situation-specific—community policing is a good example; so are family medicine and midwifery which are increasingly responsive to urban and social class differences; so is the provision of accident and emergency services and the whole development of the general hospital; national symphony orchestras and opera houses offer the same example.

This is to say that one of the 'adventures' this adolescent discipline is having at present is involvement in the reconceptualization of the role and character of the professions and their institutions for a rapidly changing democracy. There are people in police forces, hospitals, universities, the judicial system, orchestras—even jockey training schools—who need evaluative assistance to adapt to changing social contexts and new social values. I have conducted evaluations of national police training programmes in which what was sought was evaluative comment on the development by the training system of certain vocabularies, procedures and organizational roles; of bilingual education (in the USA) where what was asked for was a public shift in the policy/educational discourse; of a national symphony orchestra where the contract specified only 'to tell the story' of its schools outreach programme; and of a university where the single specified brief was to generate controversy ('light brush-fires') so as to air issues. These were all external (with the exception of the latter), independent, critical—in Scriven's terms, done 'the hard way' (1993: 64). They were all conducted more or less under the rubric of MacDonald's democratic evaluation model (MacDonald, 1977, 1987) which specifically eschews either a consensus approach or proceeding with the official definition of the programme. They all were required to involve what Scriven dismisses as 'rich description' (better to use the classic term by Geertz of 'thick description' which was never intended to refer to richness, but to an insider representation of culture), because what was required was enhanced understanding of professional culture. There was no dominant requirement in any of these to arrive at judgements of worth or merit or 'product satisfaction' ratings—although 'consumers' were consulted ... usually as citizens.

Scriven's conception of programme evaluation ('there is no need to abandon the product evaluation paradigm', 1993: 22) is not one with which I take exception in itself—I can imagine myself evaluating under his rubric. But this is a minority sport—unless we arbitrarily define programme evaluation as only that which falls inside his definition, ignoring the empirical reality of a developing professional practice. Within this empirical reality there are real demands in the professions and social institutions to extend the role and character of programme evaluation. As we have learned that business enterprises often confound classic microeconomics by prioritizing sub-optimal goals (e.g. employee satisfaction or market stability rather than profits), so we are learning that programme managers are prepared to sacrifice merit and worth (and knowledge thereof) for better cognitive control—i.e. better understanding. 'Thoroughly modern' as Scriven may be, he is, surely, behind the game.

My own prediction is that, as there is a significant time-lag in governments catching on to popular drifts, there is a significant gap opening up between government and social institutions. The Government is still hung up on amateur accountancy—products, outcomes as measures of quality-of-process, universal measurement systems of merit, and functional rationality and hierarchy as a basis for organization; while our social institutions are increasingly drawn to negotiated/responsive policy development, community accountability, situation-specific measures of merit *and* utility, flattened management, uncertainty as a condition of professional knowledge and what Eisner (1969) would call 'expressive objectives'—i.e. those objectives which emerge out of process and which are unpredictable in their outcomes. There is a tension between 'quality' and 'control'. This gap, its understanding and bridging will become major themes of programme evaluation in turbulent years ahead. Conventional (product/consumer-centred) methodologies will not serve us well.

Search for the Hero

Finally, I cannot leave this without making some comment on Scriven's use of language which is somewhat ... well ... macho. He talks of 'hard choices' (1993: 62), eschewing the 'soft-hearted route' and doing evaluation 'the hard way' (1993: 64); evaluators should 'stay mean' and be prepared to 'play rough' (1993: 87); people who will 'lie, cheat and steal in order to thwart a proposal' to be evaluated (1993: 87); evaluators may need help to 'survive their attacks' that are beyond reason; the evaluator to these people might be a 'therapist' or a 'referee' (1993: 88); evaluators should 'only want to save souls' (1993: 92); and he closes the piece with a reference to body building ... his final words are 'no pain, no gain' (1993: 92).

Here is a heroic view of the evaluator—which I again take seriously. The fascination of evaluation is precisely, as Scriven suggests, in its psychological demands. But these demands are, surely, for self-knowledge, not for iron-plated resolve. I often encounter conflict in my evaluation (no more so than any other, I hasten to say) and my first reaction is always defence ... my second reaction is always to suppress the first. I hold all the cards anyway in an evaluation—and that is precisely the problem and the source of most conflict. What I need is the self-discipline not to use them and that can demand 'high levels of' self-perception—which is the only route to diminishing the self.

I think some of the conflicts evaluators encounter require them to 'search for the hero

inside yourself ... for the secrets you hide', as the pop song goes. These conflicts surge up through the natural processes of evaluation and there is no need to provoke them. What, in my experience, certainly does provoke them is evaluators holding unsophisticated views of their and a programme's audiences (consumers). I have never read a 'thick descriptive' account of one of Scriven's evaluations but I would be surprised if, in his practice, he conformed to the implications of these vocabularies, if he held such a limited and punitive view of the citizen. I would be surprised because he is a successful practitioner as well as a leading theorist.

This is, I think, bound up with Scriven's urgency to promote evaluation as an essential social activity—his penultimate of 31 theses in the monograph is that 'evaluation is as important as content in educational programmes' (1993: 88)—again, I broadly agree. Since this is a controversial proposition we need robust people to do a robust job. Scriven's mistake is to confuse the essentiality of evaluation with the essentiality of evaluators. His (1996b: 161) list of 'minimal' (!) evaluator-competencies is formidable and certainly would exclude me and all my own colleagues (including, I am relieved to say, Barry MacDonald) from enjoying that status. In a common slip, Scriven moves from specification of need to specification of technology—and then translates that technology into correspondingly specified roles and categories of action. There are alternatives to such a technicist view.

The ideal 'programme' world (receding as fast as Earth from the Galileo space probe) is one which sports an evaluative *culture*. Here is where consumers can finally relax from Scriven's harsh demands that they deny their interest in the goals of social programmes and they can ease themselves into informed, autonomous judgement and dispense with the services of the fussy evaluator. This is where my kind of evaluation and Scriven's depart, for I would happily yield my professional autonomy in such a culture and set out with renewed curiosity and vigour to discover and document heroism in others. I suspect Michael would not so yield because, for him, consumers cannot be heroes.

Endnote

The following is the text of an email message sent to me by Michael Scriven in response to a draft of this paper.

> Kushner says, 'Michael would not so yield because for him consumers cannot be heroes', but Michael says, 'consumers are the only heroes'.
>
> Witness the struggle over the emergence of consumers' unions, the Thalidomide fight, the Vietnam war (the draftees as consumers with respect to the military programme) tec. I'm just a front man for consumers. You're right to point out their frequent interest in more than the outcomes. Often they have an interest in process (as with parents and their children's schools). Their intuitions are pretty closely matched in the checklist I use for full evaluation (Process, Outcome, Cost, Comparisons, Generalizability). The complications just make the task of the evaluator who cares for their interest more complicated; they must at all times be addressed seriously. Often, however, the consumers don't give a flying frisbee for anything but the bottom line; and in that mode which I think is the most important first line of attack (macho lives), the machinery is typically of no importance.

References

Adelman, C. (1996) 'Anything Goes: Evaluation and Relativism', in *Evaluation* 2(3): 291–305.

Cronbach, L. J., S. R. Ambron, S. M. Dornbusch, R. D. Hess, R. C. Hornik, D. C. Phillips, D. F. Walker and S. S. Weiner (1980) *Toward Reform of Programme Evaluation: Aims, Methods and Institutional Arrangements*. San Francisco: Jossey-Bass.

Eisner, E. W. (1969) 'Instructional and Expressive Educational Objectives: Their Formulation and Use in Curriculum', in W. J. Popham, E. W. Eisner, H. J. Sullivan and L. L. Tyler (eds) *Instructional Objectives*, AERA Monograph Series on Curriculum Evaluation No. 3. Chicago, IL: Rand McNally.

Ericson, D. P. (1990) 'Social Justice, Evaluation and the Educational System', in K. A. Sirotnik (ed.) *Evaluation and Social Justice: Issues in Public Education, New Directions for program Evaluation No. 45*. Oxford: Jossey-Bass.

Guba, E. and Y. Lincoln (1989) *Fourth Generation Evaluation*. London: Sage.

House, E. R. (1993) *Professional Evaluation: Social Impact and Political Consequences*. London: Sage.

House, E. R. (1995) 'Putting Things Together Coherently: Logic And Justice', in D. M. Fournier (ed.) *New Directions for program Evaluation No. 68*. San Francisco: Jossey-Bass.

Kushner, S. I. (1996) 'The Limits of Constructionism', in *Evaluation* 1(2): 189–200.

Kushner, S. I. and A. Marfleet (1995) *The Space Between Schools*. Evaluation report of the Common Course Structure at the University of East Anglia. Norwich: CARE, University of East Anglia.

MacDonald, B. (1977) 'A Political Classification of Evaluation Studies', in D. Hamilton, D. Jenkins, C. King, B. MacDonald and M. Parlett (eds) *Beyond the Numbers Game*. London: Macmillan

MacDonald, B. (1987) 'Evaluation and the Control of Education', in R. Murphy and H. Torrance (eds) *Evaluating Education: Issues and Methods*. London: Harper Row.

Phillips, D. C. (1990) 'Subjectivity and Objectivity: an Objective Enquiry', in E. W. Eisner and A. Peshkin (eds) *Qualitative Enquiry in Education: the Continuing Debate*. London: Teachers College Press.

Schon, D. (1971) *Beyond the Stable State*. London: W. W. Norton.

Scriven, M. (1993) *Hard Won Lessons in Program Evaluation, New Directions for Program Evaluation No. 58*. San Francisco: Jossey-Bass.

Scriven, M. (1996a) 'The Theory Behind Practical Evaluation', in *Evaluation* 2(4): 393–404.

Scriven, M. (1996b) 'Types of Evaluation and Types of Evaluator', in *Evaluation Practice* 17(2): 151–61.

SAVILLE KUSHNER is Senior Lecturer in Evaluation at the Centre for Applied Research in Education at the University of East Anglia, Norwich, Norfolk NR4 7TJ, UK.