Article

The uncritical realism of realist evaluation

Sam Porter
Queen's University Belfast, Northern Ireland

Abstract
This article is a response to Ray Pawson's critique of critical realism, the philosophy of science elaborated by Roy Bhaskar. I argue with Pawson's interpretation of critical realism's positions on both natural and social science and his charges concerning its totalizing ontology, its arrogant epistemology and its naive methodology. The differences between critical realism and realist evaluation are not as significant as Pawson contends. The main differences between the two realisms lie in their approaches to the relationship between social structures and human agency, and between facts and values. I argue that evaluation scientists need to clearly distinguish structure and agency. They should also make their values explicit. The uncritical approach of realist evaluation, combined with its underplaying of the importance of agency, leaves it open to implication in the abuses of bureaucratic instrumentalism.

Introduction

This article is a response to a short but significant section of Ray Pawson's realist manifesto, *The Science of Evaluation* (2013). One of the purposes of his manifesto is to distinguish realist evaluation from rival perspectives. As part of that process, Pawson engages in a denunciation of the philosophical position from which realist evaluation was spawned – critical realism, as elaborated by the philosopher Roy Bhaskar.[1]

Pawson takes us through his reading of critical realism in seven stages, each stage being prefaced by a proposition that he ascribes to critical realism. He finds himself at one with Bhaskar for the first three propositions but parts ways with him on the fourth and becomes increasingly antithetical until, by the end of his discussion, he feels able to state that critical realism 'is a strategy with no use whatsoever in applied social enquiry' (Pawson, 2013: 71).

Corresponding author:
Sam Porter, School of Nursing and Midwifery, Queen's University Belfast, Belfast BT7 1NN, Northern Ireland.
Email: s.porter@qub.ac.uk

The form of argument used here mirrors that of Pawson's, in that I quote each of the putative critical realist propositions in turn and then discuss Pawson's commentary on them. By these means I hope to demonstrate that the weaknesses that Pawson ascribes to critical realism are for the most part unfounded, and that its differences with realist evaluation are not as crucial as he makes them out to be. Two major differences relate to the interpretation of the relationship between social structure and agency, and the role of values in evaluation research. I will argue that on both these issues, the current realist evaluation position is weak, and discuss the methodological consequences of taking a critical realist approach to them.

Propositions 1−3

The first three propositions that Pawson ascribes to critical realism, which he sees as consonant with the tenets of realist evaluation, can be summarized as follows:

- Proposition 1 (2013: 61): Physical and social systems are ordered but infinitely complex, so that theory is required to guide enquiry.
- Proposition 2 (2013: 61): Physical science and social science share the same explanatory structure in that both attempt to explain events with reference to persisting, underlying mechanisms.
- Proposition 3 (2013: 61): Physical science and social science are different in that societies, being shaped by human volition, are in a permanent state of self-transformation.

Proposition 4

This difference [Proposition 3] goes on to determine the nature of and potential for empirical inquiry in the two domains. Laboratory experiments enable the researcher to neutralise external forces and so isolate conditions in which the laws of nature are revealed. Social science cannot contrive closed systems and empirical observation alone will only produce endless, partial descriptions of an ever-changing social world. (2013: 61)

Pawson identifies Proposition 4 as the juncture where the two realisms part company. He begins by taking Bhaskar to task for the limitations he places on human sciences, using the following quotation as evidence:

For in the absence of spontaneously occurring, and given the impossibility of artificially creating, closed systems, the human sciences must confront the problem of the direct scientific study of phenomena that only ever manifest themselves in open systems – for which the orthodox philosophy of science, with its tacit presupposition of closure, is literally useless. (Bhaskar, 1979: 27, cited in Pawson, 2013: 64)

Because Pawson seems to be implying that this statement relates to Bhaskar's view of the uselessness of social science *in toto*, it is necessary to clarify that Bhaskar is not referring to the consequences of openness for social science in general, but for positivism ('the orthodox philosophy of science') in particular because of its erroneous reduction of laws to empirical regularities. The irony is that, in his own critique of a variant of this orthodoxy, the 'augmented trials perspective', which seeks to establish empirical regularities in social systems, Pawson comes to the very similar conclusion that 'the whole thrust of decades of debate on

complexity is that the uniformity demanded in this scenario cannot be realised in social interventions' (2013: 49).

For Pawson, Bhaskar's limited ambitions for the social sciences are matched by his over-ambitious conception of the natural sciences. He makes a number of important claims about errors in the critical realist view of science that can be summarized under three headings:

(i) Bhaskar's model of experimentation is erroneous because experiments do not, as he claims they do, involve absolute closure

Pawson asserts that in Bhaskar's model of experimentation, 'complexity is managed by cleansing from the investigatory apparatus all other confounding influence thus allowing the physical law to unfold' (2013: 64). The crucial word in this sentence is 'all' in that Pawson contrasts what he sees as Bhaskar's model of total closure with his own model of partial closure. He makes this charge on the basis of a quotation of Bhaskar's description of experimentation as: 'an attempt to trigger or unleash a single mechanism or process in relative isolation, free from the interfering flux of the open world' (Bhaskar, 1986: 35, cited in Pawson, 2013: 64). It is immediately clear from Bhaskar's qualifications that Pawson's absolutist reading of him is unfounded. The words 'attempt' and 'relative' are there in black and white and are there for good reason, as we shall see below.

(ii) Bhaskar cannot take account of the contingent nature and incremental progression of empirical scientific endeavour

Pawson argues that critical realism's fundamentalist position on the issue of absolute closure reflects an unrealistic view of the way science works in the real world:

> In an infinitely complex world there are countless forces to be reckoned with. Science cannot know them all in advance or it would already be complete and we would have an entire state description of all the laws of physics. So the fallible scientist has to rely on some guesswork, some judgement, and a developing, collective understanding of what seems to be important in these 'sorts of systems'. In other words the controls put in place are based on hypotheses. (2013: 64–5)

Quite so, or as Bhaskar puts it:

> Science . . . must be conceived as an ongoing process of transformation, continually or essentially in motion, in an attempt to capture . . . the stratification of the world . . . Science operates on given materials, including pre-existing theory and antecedently established facts, with given materials, i.e. by means of an ensemble of intellectual and technical tools . . . producing new theories and facts. Science is produced by the imaginative and disciplined work of men [sic] on what is given to them. (2008: 185)

Far from regarding experimentation as an all-encompassing process, capable of shutting out the entire universe, Bhaskar is clear that it is always a matter of human labour, with all the compromises that such labour entails:

> The weakness of previous analyses of experimental activity is that they have not appreciated the significance of the fact that conjunctions of phenomena have to be worked for practically (as well as

in thought); that conjunctions are not given to, but *made* by us. (Bhaskar, 2008: 54; emphasis in original)

Bhaskar's description of conjunctions being made rather than given is of considerable pertinence because Pawson seeks to contrast Bhaskar's model of experimentation with his own, in which scientists 'create' a system 'that enables the regularity to unfold' (2013: 65). This dichotomy cannot stand given that both Bhaskar and Pawson accept that human creativity lies at the heart of science.

In order to illustrate the unrealistic nature of Bhaskar's account, Pawson concludes this part of his argument by using Bhaskar's own examples from experimental physics to demonstrate that even these do not involve total closure. This exercise is rendered redundant if we accept Bhaskar's description of experimentation as 'an attempt' to attain 'relative isolation' (1986: 35). Moreover, by fixing on Bhaskar's examples, Pawson misses the opportunity to uncover the most significant weakness in Bhaskar's model of science, namely his extrapolation from experimental physics and chemistry in particular to natural science in general, thus ignoring 'a range of natural sciences in which experimental closure is not an available means of empirical control on theory' (Benton, 1988: 19). Bhaskar (1989a, postscript) concedes this point to Benton, explaining that because his argument was an immanent critique of positivism, he confronted it on its chosen terrain of 'fundamental' scientific procedures in order to demonstrate that naturalism[2] was possible even under these terms of discussion, but he did not intend this to be an exhaustive account of natural science, merely a start to its analysis from a realist perspective. While neither Bhaskar nor Benton regard this retrenchment as a challenge to the fundamental tenets of critical realism, as we shall see below, acceptance that the ability to attain closure is not essential to natural science has significant consequences for the possibility of naturalism.

(iii) Bhaskar erroneously argues that social science cannot attain the closure of natural science and because of this inability, it cannot get beyond description of surface events

Pawson asserts that, according to Bhaskar, 'empirical research in social science can only ever follow and describe the surface, fleeting, ever-emergent texture of history' (2013: 65). Bhaskar says nothing of the sort. Indeed, the whole point of his treatise on the possibility of naturalism (Bhaskar, 1989a)[3] is to demonstrate the explanatory promise of the application of critical realism to empirical social science. Bhaskar does indeed assert that, in contrast to natural science, the objects of social science manifest themselves in open systems, which cannot be experimentally closed. This means that 'the social sciences are denied, in principle, decisive test situations for their theories' (Bhaskar, 1989a: 45). However, he does not conclude from this that social science is unable to penetrate the surface. Rather, for Bhaskar the only epistemological consequence of the absence of closed systems is 'that criteria for the rational development and replacement of theories in social science must be *explanatory* and *non-predictive*' (1989a: 45–6; emphasis in original).

Pawson does not accept Bhaskar's limitation on the possibility of naturalism. In commending his own naturalist model, he concludes that 'there is no need for social scientists to turn their heads to alternative strategies' (2013: 67). On this issue, he stands on firmer ground. Bhaskar's argument that the social sciences are denied decisive test situations *in principle* is

not convincing. It is a matter of scale. In relation to macro-social structures, Bhaskar is correct in stating that experimental closure is impossible. However, in terms of micro-social interactions, experimentation is perfectly possible, as the outputs of social psychologists readily attest. In between lie disciplines like evaluation science, where the scale of investigation rules out experimental closure, but where it is possible to substitute other forms of testing. For example, at this meso-level of research, randomized controlled trials can be characterized as epidemiological proxies for scientific experimentation (Blackwood et al., 2010).

When we combine acceptance of the differing abilities of different social sciences to attain closure with the same sort of differentiation within the natural sciences that was noted in the previous section, it is clear from the continuum that emerges that many natural and social sciences share the same sort of methodological problems. On these grounds at least, we can conclude that Bhaskar's position involves too stark a conception of the nature/culture divide, and that Pawson's greater confidence in the possibility of naturalism is vindicated.

However, that is not to say that there are no categorical differences between natural and social sciences. While openness in nature is a matter of complexity, in society it is a matter of both complexity and choice. This has a significant inhibitory effect on the social sciences' capacity to predict outcomes. It also requires an additional level of analysis for social science, which needs to examine both reasons and causes.

In summary, Pawson is correct in observing that some forms of social science do have available to them testing procedures that enable greater possibilities of naturalism than Bhaskar allows for; however, he is incorrect in asserting that Bhaskar's position condemns social science to surface description.

Proposition 5

The fact that the laws of nature are only produced in artificial closed systems does not mean that they are in any way under human control. Laws are laws but their sphere of operations is at an underlying level, they endure whilst not acting in the open environment. (2013: 62)

Pawson's objection to critical realism here is that it attaches too much importance to causal laws (or underlying generative mechanisms, as they are termed by realists in order to underline that their powers may be latent and that their effects will be contextually contingent).[4] In order to demonstrate this, he refers to Bhaskar's typology of domains (Table 1).

Pawson's objection to this is that the 'Domain of Real' consists of mechanisms, events and experiences, from which he concludes that 'events and experiences only become real when they are subsumed under the action of their underlying mechanism. They are only real within the totalising, transcendental system' (2013: 68).

Table 1. Domains of Reality.

	Domain of Real	Domain of Actual	Domain of Empirical
Mechanisms	✓		
Events	✓	✓	
Experiences	✓	✓	✓

Source: Bhaskar (2008: 13).

I hope to show that this conclusion entails a significant misreading of Bhaskar's typology. Essentially, the typology is a schema to show how different philosophical perspectives answer a series of questions about what exists. The first question asks 'Is reality confined to the things and events that we experience?' Empiricists would answer 'yes, *esse est percipi*'. However, actualists and realists would argue that things can exist and events can occur independently of our knowledge of them. Because there is more to reality than that which is experienced, the domain of the empirical involves an incomplete conception of reality. This leads to the further question: 'Is reality confined to things and events both observed and unobserved'? Here, actualists would answer 'yes, all there are, are things and events, and what we term laws are merely the constant conjunctions of those events'. However, realists would argue that causes are not the same as effects. Because we need to take account of the underlying mechanisms that generate events, the domain of the actual, while wider in scope than the domain of the empirical, still involves an incomplete conception of reality. The final question asks 'Is reality confined to things and events both observed and unobserved, and to the mechanisms that pattern those events?' Realists would answer 'yes' to this question. We have arrived at the domain of the real.

Bhaskar's typology is simply a description of empiricist, actualist and realist ontologies, which demonstrates the comprehensiveness of realism in comparison to the other two philosophical positions, an understanding that Pawson has already accepted. However, Pawson sees something far more sinister going on and accuses Bhaskar of asserting that 'only under total laboratory control are experiences and events brought into line with reality', and on the basis of this accusation decries his typology as 'preposterous' because 'we cannot wait upon experimental closure to reveal transcendental reality because there is no such thing as the closed experiment'. To substantiate his argument, he returns to his account of incremental physics experiments, concluding that 'it is perfect nonsense to declare that this intricate set of conditions and contingencies and of events and experiences are not yet *real*' (2013: 69; emphasis in the original). Of course it is perfect nonsense, but Bhaskar makes no such declaration. Bhaskar's typology of domains is about what exists and has nothing to say about how we can know about what exists. Given that experimentation is a methodological strategy to gain knowledge, Pawson's talk of experiments and how they can or cannot establish reality is beside the point. By conflating ontology with methodology, he falls into one of the logical fallacies that Bhaskar repeatedly warns against, namely the epistemic fallacy that 'consists in the view that statements about being can be reduced to or analysed in terms of statements about knowledge' (Bhaskar, 2008: 36).

This confusion can be traced back to Pawson's erroneous attribution to critical realism in Proposition 5 of the belief that 'laws of nature are only produced in artificial closed systems' (2013: 62). This is the opposite of the critical realist position, which is that laws of nature can *never* be produced in closed systems, which can only produce constant conjunctions. Put another way, the powers of natural generative mechanisms (the realist descriptor for natural laws) are what they are, independent of human activity. The point of experiments is to identify those powers by producing patterns of events by means of closure. To regard mechanisms and events (or even more narrowly, experienced events) as the same thing, which one must do in order to conceive of experimentation as producing laws, is to fall into the Humean errors of actualism and empiricism that Bhaskar's typology graphically illustrates.[5]

To sum up, in putting forward his typology, Bhaskar is simply providing a description of different ontological positions. To the extent that he is making a partisan point in favour of

realism, his sentiment is inclusive not exclusive. Rather than asserting that reality '*only* shows its face' in the 'constricted domains' of 'transfactual, nomic, structural essences' (Pawson, 2013: 69), Bhaskar is asserting that reality *does not only* show its face in the constricted domains of empirical experience or actual events.

Proposition 6

> Although the social world transforms ceaselessly, this does [not?] mean that it is endlessly chaotic. Social institutions change but not into blue cheese. There are corresponding social mechanisms that govern and limit social transformations. These too are located in the deep, underlying structures of society. (2013: 62)

Pawson's main focus of attention in this section is Bhaskar's transformational model of society. He rehearses three claims that Bhaskar makes about society. First, that it is simultaneously the condition of, and reproduced by human agency. This is a position that Pawson has already described as 'impeccably correct but absolutely mundane' (2013: 63) in his discussion of Proposition 3. Second, that because human agency involves choice, specific prediction of the direction of social change is impossible. Once again, this claim reflects the agreed position on the consequences of volition in Proposition 3. Third, Bhaskar observes that choices are circumscribed by the social conditions and resources available to the social agent. Such a position is consonant with Pawson's view of the necessity for social scientists to adopt a model of generative causation, as expressed in his discussion of Proposition 2 (2013: 63).

Thus far, there seems very little to fall out over. We now get to the point where Pawson evidently sees Bhaskar going awry:

> Real choices require conditions and resources to actualise them, so what the choice-makers are doing is 'self-consciously transform [sic] the conditions of existence (social structures) so as to maximise the possibilities for the development and spontaneous exercise of their natural species powers'. And with this bound Bhaskar is free. What is real about society is the totality. (2013: 70)

I am unable to follow Pawson's argument here. I cannot see how Bhaskar's characterization in the above quotation of people as self-conscious agents with the potential capacity to change social structures to their benefit adds anything significant in terms of totalizing discourse to the agreed positions that came before it. The 'bound' into totality that Bhaskar has supposedly made eludes me.

So, what does Pawson believe this totality consists of? It consists of 'the pre-existing structure as well as the limited, emerging relations that can follow under human interpretation' (2013: 70). Once again, this description involves a misreading of Bhaskar, for whom there simply is no such totality because human agency and social structure are radically different things. He could not be more insistent on this point:

> The importance of distinguishing, in the most categorical way, between human action and the social structure will now be apparent. For the properties possessed by social forms may be very different from those possessed by the individuals upon whose activity they depend . . . I want to distinguish sharply then between the genesis of human actions, lying in the reasons, intentions and plans of human beings, on the one hand; and the structures governing the reproduction and transformation of social activities, on the other. (Bhaskar, 1989b: 79)

It is precisely because of critical realism's refusal to subsume thinking human agents into a 'slowly mutating totality' (2013: 70) that it has developed the distinct methodological approach to examining the social world that involves the discrete stages of hermeneutical interpretation and identification of social mechanisms.

Before concluding, Pawson goes on to identify two aspects of the transformational model that presumably exemplify for him Bhaskar's erroneous view of society. The first relates to the relatively enduring nature of social structures; the second involves Bhaskar's adoption of a relational conception of sociology. Because Pawson does not enlighten us as to why these aspects of Bhaskar's thought are problematic beyond accusing them of being part of 'a decidedly slippery chain of reasoning' (2013: 70), there is little that can be said in response beyond observing that both have a respectable sociological pedigree in, for example, the work of Margaret Archer (1995) in relation to the former, and of Pierre Bourdieu (1990) in relation to the latter. As such, if they are to be dismissed, it would be inappropriate to do so without considered argumentation.

To sum up, there is little basis for Pawson's accusation that for critical realism 'it is the slowly mutating totality that is real' (2013: 70).

Proposition 7

> In social sciences, experimental closure not being possible, an iterative spiral of theoretical inquiry and philosophical critique can take us a long way to explanatory adequacy. Abstraction and conceptual clarification become the analogues for experimental inquiry. The objects of social science inquiries include ideas. Soundly based theoretical ideas may be used as the basis of criticism of false beliefs (2013: 62).

Pawson's concluding discussion ranges across a number of assertions about the intellectual poverty of critical realism, which I will address under four headings.

(i) The dogmatism of critical realism

The assertion here is that critical realist social science is as dogmatic as critical realist natural science, such that the difference between them is that 'indisputable experiments are replaced by indisputable theories' (2013: 71).

Pawson's opening observation that Bhaskar claims that social science tends to require more theoretical work than natural science, 'and, moreover, it needs to come first' (2013: 70) is fair comment. What is not fair is his implication that this is inappropriate. Let us jump forward a chapter to consider Pawson's justification for putting theory as the first step in the organizing principles of his own recipe for evaluation science: 'The real starting point of science . . . lies in "theory", our ideas on the nature of the problem and on the nature of its solution' (2013: 87).

In order to demonstrate his point about Bhaskar's claims concerning the indisputability of theory irrespective of evidence, Pawson includes and comments on the following quotation:

> 'Thus in social science attempts at real definitions will in general precede rather than follow successful causal hypothesis' (Bhaskar, 1979: 63). The conquering strategy for social science thus boils down to the production and celebration of _a priori_ reasoning. (2013: 70)

The quotation chosen by Pawson is convenient for his argument. The trouble is that he has given us only the first half of Bhaskar's sentence. Let me provide the second half: 'though in

both cases they can only be justified empirically, viz. by the revealed explanatory power of the hypotheses that can be deduced from them'. It is clear from consideration of the full sentence that Pawson's accusation of anti-empirical dogmatism on the part of Bhaskar is unfounded.

Pawson is on stronger ground when it comes to the next quotation he takes from Bhaskar: 'the substantive employment of essentially apodeictic argument should occasion no surprise' (Bhaskar, 1979: 64, cited by Pawson, 2013: 70). Pawson's objection centres on the word 'apodeictic', for which he supplies two definitions: 'clearly established' or 'beyond dispute' (2013: 71).

What is Bhaskar referring to when he uses the term 'apodeictic'? As Pawson observes, he is laying claim to the possibility of synthetic a priori knowledge. Of course, this move from Greek to Latin doesn't get us much further in terms of clarity. What Bhaskar is saying is that it is possible to use logic to establish what must be the case for the features we observe in the world to be possible. Following Kant (1970), who first described it, this process is termed 'transcendental'. Bhaskar argues that, when asking the transcendental question, if our premises are correct, and if the trail of logic that we use is sound, then we can be confident of our conclusions (this is where the term 'apodeictic' comes in, though as we saw in the redacted quotation above, because of the 'ifs' embedded in this process, we still need empirical testing). Thus, for example, if we ask what must be the case for objects to tend to fall towards the earth, then the answer is that there must be a generative mechanism with the power of attraction. However, there are objects, such as birds and aeroplanes, which at least temporarily defy this law. The realist argument that events are governed by the interaction of a multiplicity of generative mechanisms can explain this, but it is my suspicion, and here I have some sympathy with Pawson, that the more complex the interaction, the less confident we can be that the use of pure logic will lead to 'apodeictic' answers. The objection becomes even stronger when it is applied to the actions of human beings. The point is at least partially conceded by Bhaskar through his acceptance that the transcendental move can only uncover necessary, not sufficient conditions for events.

However (and this is a crucial qualification), it is not the case that Bhaskar's assertion of the apodeicticity of transcendental logic means that he is adopting the arrogant infallibility that Pawson accuses him of. This is because he accepts that logic can only work on the materials that it is given:

> Both the ultimate premises and the immediate conclusions of philosophical considerations are contingent facts, the former (but not the latter) being necessarily social and so historical. It is only in this relative or conditional sense that philosophy can establish synthetic a priori truths (truths about the world investigated by science). Philosophy operates by pure reason. But it does not operate by the use of pure reason alone. For it exercises that reason always on the basis of prior conceptualisations of historical practice, of some more or less determinate social form. (Bhaskar, 2008: 259)

So, even if we do not share Bhaskar's confidence in the capacity of logic, that does not mean that we are warranted in accusing his philosophy of science of positing the possibility of 'indisputable theories' (Pawson, 2013: 71), given that Bhaskar accepts that theories are always socially grounded and therefore contingent.

(ii) The Marxism of critical realism

As a substantive example of the presumptuousness of transcendental realism, Pawson points us to Bhaskar's commentary on the realism of Marx's analysis of capitalism, quoting Bhaskar's

observation that 'Capital [the book] may most plausibly be viewed as an attempt to establish what must be the case for the experiences grasped by the phenomenological forms of capitalist life to be possible' (Bhaskar, 1979: 65, cited by Pawson, 2013: 71). He takes this statement to be an assertion by Bhaskar that the truth about the nature of economic relations has been firmly established by Marxist analysis, so that 'we already know that the real causes lie in the institutions and market forces of the capitalist system' (2013: 71). Pawson's failure to take cognisance of Bhaskar's qualifications is again the Achilles' heel of his argument. Even though he has written it out verbatim, he ignores the two important qualifications contained in the phrase 'may most plausibly be viewed as an attempt'. First, Bhaskar rests the adequacy of his own interpretation of Marx on plausibility rather than certainty, and second, he describes Marx's effort as nothing more than an attempt, adjudication of the substantive success or failure of which is beyond the remit of an underlabouring philosophy[6] like critical realism, the purpose of which is to concentrate on the processes of science, not its outcomes. In short, Pawson's assertion of dogmatism on the part of critical realism is not supported by the evidence quoted.

Pawson also uses Marxism to introduce another variant of critical realist totalization. To illustrate critical realism's attachment to indisputable theory, he quotes Bhaskar's statement that 'it has always been the special *claim* of Marxism to be able to grasp social life as a totality' (Bhaskar, 1979: 65, cited Pawson, 2013: 71, my emphasis). Bhaskar's sentence continues (unquoted): 'to display it, in Labriola's words, as "a connection and complexus"'. Bhaskar accepts this particular notion of totality – that specific aspects of the social world connect with each other in a matrix of relationships so that no social form exists in isolation, sufficient unto itself. The consequence that this has for social scientists is that, depending on the question they are asking, they may need to look beyond the particular towards the matrix for their answer. As far as evaluation science is concerned, while the term 'totalization' may ring alarm bells for Pawson, such an approach is everyday good practice. Thus, for example, Greenhalgh et al.'s (2004) specification of the 'outer context' as one of the key components that health services evaluators need to address involves just such an assumption.

Pawson is correct in attributing to Bhaskar an adherence to totality in the sense that Bhaskar views the social world as an interconnected 'complexus'. However, the fact that this conception is commonplace in (and, arguably, essential to) social science neutralizes his critical intent. Moreover, he is incorrect in attributing to Bhaskar acceptance of Marxist claims to be able to 'grasp' and 'display' that complexus. The obvious clue is in Bhaskar's qualification – the word 'claim'. Bhaskar goes on to use the term again: 'Marxism has claimed to be able to do so in virtue of a theory of history . . . A theory of history can only be judged by historical materials' (Bhaskar, 1989a: 43–4), hammering home this agnosticism by noting that critical realism allows for comment on the intentions but not on the results of the Marxist project.

(iii) The arbitrary values of critical realism

In addressing the 'critical' component of critical realism, Pawson describes it as a strategy for dealing with 'the relentless play of human choice . . . by deciding *a priori* which choices are truly human and which bear the marks of false consciousness' (2013: 71). While he does not spell out what he means by 'truly human', it seems to refer to some kind of essential human authenticity that can be used to judge the choices people make. It appears that Pawson is pinning the concept of true humanity on Marxism, and by extension on Bhaskar. Perhaps he has

in mind the Marxist notion of 'species-being', given his earlier quotation of Bhaskar's term 'natural species being'. This refers to the position that, like all other living species, humans, by dint of their evolutionary status, have certain capacities. Some, such as the capacity to acquire nutrition, they have in common with all living species. Some, such as the capacity to self-propel, they have in common with most animal species. Some, such as the capacity to make rationally-guided decisions (*vide* Kant, 1970); to formulate normative judgements (*vide* Sayer, 2011); to engage in sophisticated symbolic communication (*vide* Habermas, 1981); or to creatively manipulate their environment (*vide* Marx and Engels, 1970) are unique to the human species. Thus, rather than being based on some kind of essentializing moral claim, the notion of species-being is grounded in evolutionary biology, and hence is amenable to empirical interrogation.

While not specifically addressed by Bhaskar, false consciousness is a more apposite target, given that Bhaskar (1989a) discusses the concept of ideology using Marx's critique of political economy as an example. Ideology is the classical Marxist description for a system of false beliefs, so while not exactly the same as false consciousness (which is its phenomenological instantiation), it is a closely related concept. If it can be demonstrated that Bhaskar's procedure for identifying ideology involves the application of a priori criteria, then it is reasonable to extrapolate that the identification of false consciousness would be similarly arbitrary.

The first criterion given by Bhaskar (1989a) for the identification of ideology involves demonstrating a theory's comparative explanatory weakness. This entails assessing whether a rival theory can explain most or all of the phenomena explained by it, and in addition explain significant phenomena not explained by it. This is essentially the application of Kuhnian (1962) precepts about the epistemological adequacy of paradigms. However, a suboptimal theoretical position can only be designated as ideological on the basis of the additional criterion that it must be demonstrated to mask the nature of the structured relationships to which it pertains. If we allow for the possibility of being able to scientifically identify structured relationships, then such a critical comparison should not pose any insurmountable methodological problems. It certainly does not require the a priori application of the notion of true humanity or any other extraneous normative standard. As Bhaskar observes, the critical process involved in uncovering ideological theories takes place 'without the intervention of any value judgements (other than those bound up in the assessment of the cognitive adequacy of the theory)', noting that Marx's critical analysis of political economy 'is subversive in virtue of its cognitive power alone' (1989a: 70).

It is ironic that Pawson, in accusing critical realism of both explanatory inadequacy and of masking politics with scientific rhetoric, faithfully adopts Bhaskar's criteria for the identification of ideology, thereby implying that those who espouse this 'ideology' are displaying false consciousness.

(iv) The uselessness of critical realism

Pawson's final critique of critical realism is that:

> It is a strategy for lording over complexity rather than analysing it. It is a strategy with no use whatsoever in applied social inquiry. The science of evaluation starts by recognising that the fate of social policy lies in the real choices of choice makers and its task is to explain the distribution and consequences of those choices rather than to condemn them. (2013: 71)

I wish to highlight two issues contained in this quotation. The first is the accusation that critical realism 'is a strategy for lording over complexity rather than analysing it'. In response, I wish to return to Bhaskar's insistence on the categorical distinction between structure and agency. This analytical dualism has been the basis for important advances in the understanding of social complexity. In contrast to the 'downwards conflation' of structuralism, the 'upwards conflation' of methodological individualism, and the 'central conflation' of structuration theory (all of which display, in their own ways, a totalizing approach to society), Margaret Archer's (1995) critical realist sociology allows for greater insight into the complex relationship between structure and agency through its acknowledgement that they are analytically separable and temporally sequenced.

Second, the final phrase in Pawson's closing statement contains two telling words – 'explain' and 'condemn', which take us to the crux of the axiological disagreement between realist evaluation and critical realism. They encapsulate the fact/value distinction that Pawson believes essential if evaluation science is not to 'abandon analysis for ideology' (2013: 81), the fundamental error of critical realism.[7]

In response, the first thing to note is the responsibility that arises from the fact that the discourse of evaluation science often has real consequences for people. Given that some of the most important issues that people face are normative; that they are 'beings whose relation to the world is one of concern' (Sayer, 2011: 2), there is an onus on social scientists to acknowledge those value-laden concerns. The second thing is that, as Pawson rightly points out, the social world facing the evaluator is infinitely complex. This means that the evaluator has to make judgements about, *inter alia*, the selection of issues, the emphasis placed on different factors, and the descriptions they attach to them. While these decisions are not entirely matters of value judgement, they cannot be made without recourse to values, whether through conscious consideration or unconscious acceptance. For the evaluator to hide behind the fig leaf of disinterested science is false modesty. As Turner has observed in relation to the applied social science of criminology:

> As we now are more aware than ever that what emerges as 'knowledge' at any given time, in any given place, is contingent upon the context within which such knowledge is produced, it will not do to deny the choices that the social scientist inevitably makes . . . To misrepresent the form and content of social scientific knowledge as somehow wholly mandated by the empirically observable and objective 'facts' of reality . . . is to make social scientists into Gouldner's (1970: 13) 'moral cretins'. (2013: 162−3)

If values are inevitably present in the decisions of the social scientist, then it is the responsibility of that scientist to consider carefully which values to adopt to guide decision making, and to be transparent about those values. This is all the more important when the science concerned plays a significant public role, at a time when concerns are increasingly being raised about the deindividuating effects of evidence-based policy and practice (Porter and O'Halloran, 2009). Realist evaluation's lack of robust critical values and adherence to Popperian (1945) piecemeal social engineering (Tilley, 2000) pushes its focus onto technical solutions and away from 'ideological' concerns such as whether the solutions it posits promote or inhibit non-alienating social forms. From here, it is but a short step to a technocratic consciousness (Habermas, 1971) that transmutes personal and ethical problems into scientific and technological categories amenable to purposive-rationality, and in the process devalues the lifeworlds

of those whom it affects (Porter and O'Halloran, 2012). In contrast, critical realism's valorization of transformative social action 'to maximise the possibilities for the development and spontaneous exercise of [people's] natural (species) powers' (Bhaskar, 1989a: 37), while open to debate, amendment and elaboration, provides a solid guard against the slide into technocratic consciousness.

To summarize, I have argued that Pawson's attempt to characterize critical realism as claiming the indisputability of theory is contradicted by Bhaskar's repeated assertions about the need for empirical confirmation. Pawson's attempts to impute a substantive Marxist position to critical realism are also unfounded. His accusation that critical realism lords it over complexity fails to recognize its contribution to understanding the complex relationship between social structure and agency. Finally, the claim that evaluation science should eschew the critical use of values involves a derogation of the responsibility to address people's concerns.

Methodological consequences

The arguments presented thus far have, by necessity, been somewhat esoteric. I now wish to establish their pertinence to the practices of evaluation science by identifying some of their methodological consequences. Specifically, I will discuss the outworking of two critical realist positions that distinguish it from realist evaluation: its acceptance of the categorical distinction between social structure and agency; and its rejection of the absolute separation of facts and values.

Differentiating structure and agency

In Proposition 6 Pawson accuses Bhaskar of subsuming human agency into a structurally-dominated totality. I refuted this charge by recalling Bhaskar's distinction 'between the genesis of human actions, lying in the reasons, intentions and plans of human beings, on the one hand; and the structures governing the reproduction and transformation of social activities, on the other' (Bhaskar, 1989b: 79). I wish to push the argument a stage further here by contending that it is Pawson, in his conceptualization of programme mechanisms, who is guilty of positing a totality that consists of 'the pre-existing structure as well as the limited, emerging relations that can follow under human interpretation' (Pawson, 2013: 70).

Pawson regards social mechanisms as a combination of agency and structure, which 'are thus about people's choices *and* the capacities they derive from group membership' (Pawson and Tilley, 1997: 66; my emphasis). In contrast, critical realism conceives of social mechanisms as generative powers that result from the distribution of material, cultural and jurisdictional resources embedded in relatively enduring social relations. In other words, they sit exclusively within the structural component of the social world. Agents, while they both influence and are influenced by these social systems, are different, not least because they have the capacity to think and choose while social structures do not. This is not to deny the existence of generative mechanisms at an individual level, but the realist conception of stratification distinguishes psychic from social mechanisms. Thus, while social mechanisms will undoubtedly impinge upon people's choices, they are not the mechanisms involved in the making of choices, which reside in the psyches of individuals.

Programme mechanisms are a sub-set of social mechanisms, in that that they have been consciously created by agents with the aim of altering the interpretations and actions of other

agents. However, this still does not mean that they themselves have the characteristics of agency. Rather than entailing 'the interplay between social resources and participants' reasoning' (Pawson and Tilley, 1997: 75), programme mechanisms are only one part of that interplay, involving the structural distribution of social resources. Participants' reasoning, while taking into account the perceived consequences of that distribution, is categorically distinct.

The benefit of critical realism's dualist model is that it rescues agency from the epiphenomenal position that realist evaluation's 'context + mechanism = outcome' (CMO) formula places it in. Realist evaluation's conflation of agency and structure under the rubric of 'mechanism' means that it cannot sufficiently recognize the capacity of humans to choose, and is thus an inadequate description of causal chains in the social world. In order to adequately reflect the distinction between social structure and human agency, I propose that agency should be included as a separate category alongside context and mechanism in the evaluation formula. This addition to the formula points to the need for evaluators to adopt three distinct methodological approaches: one designed to enumerate outcome patterns; one designed to identify the mechanisms embedded in an intervention and its social context; and one designed to uncover the experiences, interpretations and responses of the actors involved.[8] This revision also has the benefit of clarifying the evaluation process. Acceptance that structure and action are analytically separate may help resolve some of the methodological confusion that has dogged evaluators attempting to operationalize the CMO formula in their work (*vide* Marchal et al., 2012).

Rehabilitating utopianism

The aspirations of critical realism have a ring of utopianism to them. By contrast, realist evaluation prides itself in being grounded, adhering to the view that science should service modest 'piecemeal social engineering' in contrast to the dangerous fanaticism of 'utopian social engineering' (Tilley, 2000), as these alternatives are set out in Popper's (1945) liberal polemic. If Popper's dichotomy is accepted, there are good reasons for evaluators to concur with this choice, not least because the programmes examined by evaluation science are almost by definition piecemeal, in that they aspire to change specific aspects of social life, rather than society as a whole.

However, if one does not accept Popper's dichotomy, it is possible for the evaluator to adopt a critical stance without falling into the trap of ends-over-means grandiosity. One approach that I have previously identified (Porter and O'Halloran, 2012) is Ernst Bloch's (1995) 'concrete utopianism', which eschews millenarianism for a combination of the cool appraisal of the 'limited possible' (the piecemeal) and an exploration of the possibilities that lie beyond this limiting horizon (the utopian), with the aim of promoting human capacities and potentialities without recourse to authoritarian tactics.

Referring back to the capacities and potentialities identified in the discussion on species-being in Proposition 7, their development and exercise requires the satisfaction of both non-species-specific needs such as adequate nourishment, and species-specific needs such as individual autonomy (Kant); creativity (Marx and Engels); and effective discursive participation (Habermas) that takes account of normative concerns (Sayer, 2011). The satisfaction or otherwise of these needs can be used as ethical criteria for judging interventions.

The data required to inform ethical judgements need to reflect agents' experience. That experience could be interrogated with a view to establishing the degree to which an

intervention supports the development of people's capacities and potentialities, and the degree to which it inhibits them. Depending on the precise nature of a programme, the phenomeno-logical experiences of those who are its putative beneficiaries; those whose behaviour it is intended to alter; and/or those who are expected to implement it should be at the core of the evaluative exercise. An intervention could then be judged on the basis of whether these actors interpret it as promoting their human needs, or whether they regard it as alienating, along with an analysis of the conditions that ground their interpretations. These proposals, with their emphasis on the centrality of human interpretation, reinforce the need for the inclusion of agency as a distinct analytic category in the evaluation formula.

This methodological recommendation is consonant with realist evaluation's mantra that evaluation research should entail the examination of 'what works for whom in what circumstances' (Pawson, 2003: 474). However, it is based on the assumption that these questions cannot be adequately answered by taking the technical approach adopted by realist evaluation, but need to be approached from an ethical standpoint that takes account of people's concerns. It focuses clearly on the experience of an intervention, rather than merely concentrating on the processes leading to its effectiveness or otherwise. Thus, it again mandates three stages in the process of evaluation: enumeration of behavioural outcome patterns; examination of the resources and restrictions embedded in an intervention and its social context; and critical hermeneutical investigation of the experiential consequences of these structural circumstances in terms of the satisfaction of human needs.

Conclusion

The main thrust of this article has been to question Pawson's characterization of Bhaskar's critical realism as entailing a totalizing conception of science.[9] I have argued that many of the putative differences between the two realisms that he posits are illusional. Moreover, I have argued that where differences do exist, there is good reason for evaluators to adopt the positions of critical realism, rather than those put forward in realist evaluation. Specifically, I have contended that evaluation of interventions needs to focus on both the social mechanisms they entail, and the responses to these by the actors affected by them. This in turn allows for the inclusion of critical hermeneutics in the process of evaluation research, with the purpose of establishing the degree to which social agents experience interventions as promoting or inhibiting the development and exercise of their natural, species powers. On this basis, interventions can and should be condemned (or praised).

If these conclusions are accepted, it presents realist evaluators with a significant choice. If they decide that there is no place for critical evaluation in their scientific model, then the challenge facing them is to demonstrate that they can resist the lure of technocratic engineering to facilitate smooth systemic running, or, if not, that they can justify the role of organic intellectual servants to bureaucratic masters as a fitting one for evaluators. If they decide that there is a place for critical evaluation, then they need be less concerned that this will require them to abandon scientific rigour in favour of 'a parody of science' (2013: 71).

Funding

Notes

1. Pawson concentrates his critique almost entirely on Bhaskar's work, and I follow him in this. However, this should not be taken to imply that I do not recognize that critical realism has been developed by numerous theoreticians and researchers in addition to Bhaskar. Nor should it be taken to imply any expectation that Bhaskar's work should be engaged with by evaluators in a catechistic fashion.
2. 'Naturalism', in the sense used here, refers to the use of natural scientific methods to study the social world.
3. While Pawson uses the first edition of Bhaskar's *The Possibility of Naturalism* (1979), my references are to the second edition (1989a).
4. 'Causal laws must be analysed as tendencies of things, which may be possessed unexercised and exercised unrealized, just as they may of course be realized unperceived (or undetected) by people. Thus in citing a law one is referring to the transfactual activity of mechanisms, that is, to their activity as such, not making a claim about the actual outcome (which will in general be co-determined by the activity of other mechanisms)' (Bhaskar, 1989a: 9−10).
5. For Bhaskar, the assumption that experiments can produce laws is the *reductio ad absurdum* of the Humean conception of causality: 'A *real* distinction between the objects of experimental investigation, such as causal laws, and patterns of events is . . . a condition of the intelligibility of experimental activity . . . [T]he Humean account depends upon a misidentification of causal laws with their empirical grounds. Notice that as human activity is in general necessary for constant conjunctions, if one identifies causal laws with them then one is logically committed to the absurdity that human beings, in their experimental activity, cause and even change the laws of nature!' (1989a: 9; emphasis in original).
6. 'Underlabouring philosophy' is a term for his approach that Bhaskar borrows from John Locke, who makes the following modest justification for his intellectual efforts: 'It is ambition enough to be employed as an under-labourer in clearing the ground a little, and removing some of the rubbish that lies in the way to knowledge' (1959: 14).
7. It might be noted that two words are a somewhat flimsy foundation on which to construct such a weighty argument. That Pawson felt able to introduce this fundamental component of his approach to evaluation science in such an abbreviated fashion is notable in itself. Fortunately, more substantial evidence of his adherence to the fact/value distinction can be found elsewhere, where he states that the fundamental division between the two realisms is 'on the matter of whether social science should primarily be a critical exercise or an empirical science' (Pawson, 2006: 18). Implicit in this statement is the assumption that it is necessary to adopt either critical values or the systematic pursuit of empirical facts. For critical realism, this is a false dilemma, in that it does not accept that the adoption of values must lead to the abandonment of facts.
8. This suggestion begs the question of how generative mechanisms should be identified and their relationship with agency established. To wit, should mechanisms be identified by means of realist synthesis and theoretical interviewing, or should they be retroduced from the experiences of actors by means of transcendental analysis? For reasons of complexity and space, such a discussion is not feasible here, beyond the suggestion that triangulation of the two approaches is a possibility that merits future consideration.
9. Following the focus of Pawson's critique, my arguments have related exclusively to the critical realist phase of Bhaskar's thought. It is the case that Bhaskar introduces totality in his middle, 'dialectical' phase. Moreover, in his later 'meta-reality' turn, he adds metaphysics to the mix.

However, these are philosophical enterprises that Bhaskar (2002) himself acknowledges as analytically distinct from critical realism, and take positions on totality and metaphysics that many critical realists do not subscribe to. I am included in their number and have no wish to defend such positions here.

References

Archer M (1995) *Realist Social Theory: The Morphogenetic Approach*. Cambridge: Cambridge University Press.

Benton T (1988) Realism and social science: some comments on Roy Bhaskar's 'The Possibility of Naturalism'. *Radical Philosophy* 27: 13–21.

Bhaskar R (1979) *The Possibility of Naturalism*. Brighton: Harvester Press.

Bhaskar R (1986) *Scientific Realism and Human Emancipation*. London: Verso.

Bhaskar R (1989a) *The Possibility of Naturalism: A Philosophical Critique of the Contemporary Human Sciences*, 2nd edn. Hemel Hempstead: Harvester Wheatsheaf.

Bhaskar R (1989b) *Reclaiming Reality: A Critical Introduction to Contemporary Philosophy*. London: Verso.

Bhaskar R (2002) *Reflections on Meta-Reality: Transcendence, Enlightenment, and Everyday Life*. Thousand Oaks, CA: SAGE.

Bhaskar R (2008) *A Realist Theory of Science*, 2nd edn. London: Verso.

Blackwood B, O'Halloran P and Porter S (2010) On the problems of mixing RCTs with qualitative research: the case of the MRC Framework for the Evaluation of Complex Healthcare Interventions. *Journal of Research in Nursing* 15: 511−21.

Bloch E (1995) *The Principle of Hope*. London: Routledge.

Bourdieu P (1990) *The Logic of Practice*. Cambridge: Polity.

Gouldner A (1970) *The Coming Crisis of Western Sociology*. New York: Basic Books.

Greenhalgh T, Robert G, Bate P, Kyriakidou O, Macfarlane F and Peacock R (2004) *How to Spread Good Ideas: A Systematic Review of the Literature on Diffusion, Dissemination and Sustainability of Innovations in Health Service Delivery and Organisation*. London: NCCSDO.

Habermas J (1971) *Towards a Rational Society*. London: Heinemann.

Habermas J (1981) *Theory of Communicative Action, Volume 1: Reason and the Rationalization of Society*. Boston, MA: Beacon Press.

Kant I (1970) *Critique of Pure Reason* (N. Kemp-Smith, trans.). London: Macmillan.

Kuhn T (1962) *The Structure of Scientific Revolutions*. Chicago, IL: Chicago University Press.

Locke J (1959) *An Essay Concerning Human Understanding*, vol. 1. New York: Dover.

Marchal B, Van Belle S, Van Olmen J, et al. (2012) Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation* 18(2): 192–212.

Marx K and Engels F (1970) *The German Ideology*. London: Lawrence and Wishart.

Pawson R (2003) Nothing as practical as a good theory. *Evaluation* 9(4): 471−90.

Pawson R (2006) *Evidence-Based Policy: A Realist Perspective*. London: SAGE.

Pawson R (2013) *The Science of Evaluation: A Realist Manifesto*. London: SAGE.

Pawson R and Tilley N (1997) *Realistic Evaluation*. London: SAGE.

Popper K (1945) *The Open Society and its Enemies: Volume 2, Hegel and Marx*. London: Routledge and Kegan Paul.

Porter S and O'Halloran P (2009) The postmodernist war on evidence-based practice. *International Journal of Nursing Studies* 46(2): 740−8.

Porter S and O'Halloran P (2012) The use and limitation of realistic evaluation as a tool for evidence-based practice: a critical realist perspective. *Nursing Inquiry* 19(1): 18−28.

Sayer A (2011) *Why Things Matter to People: Social Science, Values and Ethical Life*. Cambridge: Cambridge University Press.

Tilley N (2000) Realistic evaluation: an overview. *Presentation to the Danish Evaluation Society*. URL (consulted 25 July 2013): http://evidence-basedmanagement.com/wp-content/uploads/2011/11/nick_tilley.pdf.

Turner E (2013) Beyond 'facts' and 'values': rethinking some recent debates about the public role of criminology. *British Journal of Criminology* 53(1): 149−66.

Sam Porter is a nurse and a sociologist. He holds the Chair in Nursing Research at Queen's University Belfast. His substantive research areas include palliative care and therapeutic uses of the arts. He has written widely on social theory and methodology in relation to healthcare.