Systemic Practice and Action Research, Vol. 11, No. 1, 1998

Action Research: Its Nature and Validity

Peter Checkland[1,2] and Sue Holwell[1]

Received October 30, 1997

The process of knowledge acquisition which has the strongest truth claim is the research process of natural science, based on testing hypotheses to destruction. But the application of this process to phenomena beyond those for which it was developed, namely, the natural regularities of the physical universe, is problematical. For research into social phenomena there is increasing interest in "action research" in various forms. In this process the researcher enters a real-world situation and aims both to improve it and to acquire knowledge. This paper reviews the nature and validity of action research, arguing that its claim to validity requires a recoverable research process based upon a prior declaration of the epistemology in terms of which findings which count as knowledge will be expressed.

KEY WORDS: research; action research; research methodology.

1. INTRODUCTION

As *Systems Practice* extends its scope to include action research (AR), it is appropriate to reflect briefly on both the nature of AR and, because this is often challenged, its validity as a mode of inquiry leading to defensible and potentially transferable results. These characteristics could not of course include *replicability*, and that is the source of the challenge. Whenever sulfuric acid is added to a barium chloride solution, a white precipitate settles out. It does not matter who does it, in what part of the world, or in what circumstances; the precipitate is always formed. Chemists accept that they are investigating phenomena which are not capricious. They no longer agonize about the provenance of their discipline or the replicability of their results: they can be publicly tested. Chemists can concentrate on making substantive contributions within a taken-as-given framework.

Things are more volatile in the investigation of human and social phenomena. There it is still necessary to argue about the underlying assumptions, the

[1]Management School, Lancaster University, Lancaster LA1 4YX, UK.
[2]To whom correspondence should be addressed at 25 Pinewood Avenue, Bolton-le-Sands, Lancaster LA5 8AR, UK, or Management School, Lancaster University, Lancaster LA1 4YX, UK.

1094-429X/98/0200-0009$15.00/0

modes of inquiry, and their validity. In that context AR is a mode of inquiry particularly in need of defense. This paper, which draws on and extends Checkland and Holwell (1997, Chap. 1) examines the provenance of AR and its problems. It offers a model derived from a 25-year program of AR carried out at Lancaster University and presents an argument for an appropriate form of validation which, though it does not match the magic of the replicability criterion in natural science, can sustain AR as a legitimate form of inquiry which can be defended if it meets certain criteria which are developed here.

In what follows we are concerned with the fundamentals of "action research" and use the generic shorthand AR as covering a number of versions of the approach which have acquired such names as action learning (Revans, 1972), action science (Argyris *et al.*, 1982), action inquiry (Torbet, 1991), participatory action research (Whyte, 1991), and RAAKS (rapid appraisal of agricultural knowledge systems) (Engels and Salomon, 1997). Clark (1972) provides a useful bibliography of AR up to the early 1970s, and Dash (1997) usefully surveys this field from a more recent perspective.

An initial chronological sampling of this field is provided by Lewin (1947), Blum (1955), Foster (1972), Clark (1972), Susman and Evered (1978), Hult and Lennung (1980), Argyris *et al.* (1982), Susman (1983), and Dash (1997).

2. THE PROVENANCE OF ACTION RESEARCH

For most people who consider the matter at all, the paradigm model of organized inquiry is that provided by natural science. This is not surprising since the investigation of natural phenomena via the method of science is undoubtedly the most powerful form of knowledge generation ever devised. The development of that method is the crucial distinguishing characteristic of the civilization in which it has emerged, starting with the pre-Socratic philosophers in Ancient Greece in the 6th century BC. They postulated *rational* myths about the world, rather than myths involving supernatural beings, myths about which there cannot be much discussion. Rational myths *can* be discussed, however, and their emergence led to the development of rational methods of investigating the world, culminating eventually in the Newtonian scientific revolution of the 17th century. Our own century, with the subsuming of the Newtonian model as a limited special case of Einstein's physics, has taught us that all scientific knowledge is in fact provisional, being simply the best-tested knowledge we have at any given time. But the experimental method which generates that (provisional) knowledge is now taken as a given. This is thanks to its success, through science-based technology, in creating our worldview and our world.

The scientific method can be expressed as being based on three fundamental principles which characterize it and give it its power: reductionism, repeatability, and refutation (Checkland, 1981, Chap. 2). Scientists select a portion of the

world to investigate and carry out disciplined observations in experiments. If the results of the experiments are repeatable, they count as part of the body of knowledge; and progress can be made in sequences of experiments through the testing to destruction of hypotheses. Scientific knowledge is then the accumulation of hypotheses which have not (yet) been refuted. This method of inquiry has been so successful that, in Western culture, to declare some putative knowledge as "unscientific" is often to justify dismissing it as irrelevant.

The power of scientific method lies in the *replicability* of its results; this turns its findings into "public knowledge" (Ziman, 1968) (though argument can of course still rage concerning the *interpretation* of the demonstrable experimental happenings). This replicability of experimental results stems from the fact that the phenomena investigated must be, in Keynes' phrase "homogenous through time": the inverse square law of magnetism is always, demonstrably, an inverse square law. Keynes (1938), quoted by Moggridge (1976, p. 26), was pointing out that economics should repel attempts to turn it into a pseudo-natural science precisely because

> unlike the typical natural science the material to which economics is applied is, in too many respects, not homogenous through time.

The point which Keynes makes highlights brilliantly the difficulties for social-scientists who would like to make use of the outstanding successful method of inquiry developed in the natural sciences. Can the method of science be applied to material which is not homogeneous through time, making complete replicability impossible? If not, what else can be done?

This is the context in which AR emerged. Kurt Lewin (1890–1947), a psychologist who became interested in human groups and their dynamics, particularly from the point of view of bringing about change in society, came to perceive "the limitations of studying complex real social events in a laboratory, the artificiality of splitting out single behavioural elements from an integrated system" (Foster, 1972). The concept emerged of a researcher immersing himself or herself in a human situation and following it along whatever path it takes as it unfolds through time. This means that the only certain object of research becomes the change process itself. This is a difficult concept for those anxious to import hypothesis-testing into social research, though it is an approach with which anthropologists and sociologists are familiar. Whyte (1991, p. 9) from the 1940s, was doing work in which "informants" in situations he researched became "active participants in the research," thus blurring the distinction between the researcher and those researched. This is something which worries natural scientists and those who would emulate their method of inquiry. As Vickers used to point out (Checkland, and Holwell, 1997, p. 19), since social phenomena are mental abstractions at a meta-level to their manifestations, even *thinking and arguing* about them can change them! On the other hand, whether

the structure of our part of the cosmos corresponds to Ptolemy's earth-centered model or Copernicus' heliocentric model is entirely unaffected by our having theories about it. But Marx's theory of history changes history! Social phenomena are not, in Keynes' phrase, "homogeneous through time": hence the idea of taking part in change in organizations as a basis for research in the social world.

Probably most "interpretive" action researchers, acting on the assumption that social reality is continuously being created and recreated in a social process, would accept the notion of Argyris *et al.* (1982) that the crucial elements in a research approach which works within a specific social situation are

- a collaborative process between researchers and people in the situation,
- a process of critical inquiry,
- a focus on social practice, and
- a deliberate process of reflective learning.

This implies a very different kind of research from the testing of hypotheses. The latter process is represented in Fig. 1. This shows an "ideal-type" model of positivist research in which a researcher propounds a hypothesis about some part of perceived reality and then tries to test that hypothesis to destruction. It is useful to develop an alternative which covers the kind of approach Argyris and his colleagues describe, in order both to organize this kind of research and to explore the difficulties which AR faces, not least the question of its validity. This is done in the next section.

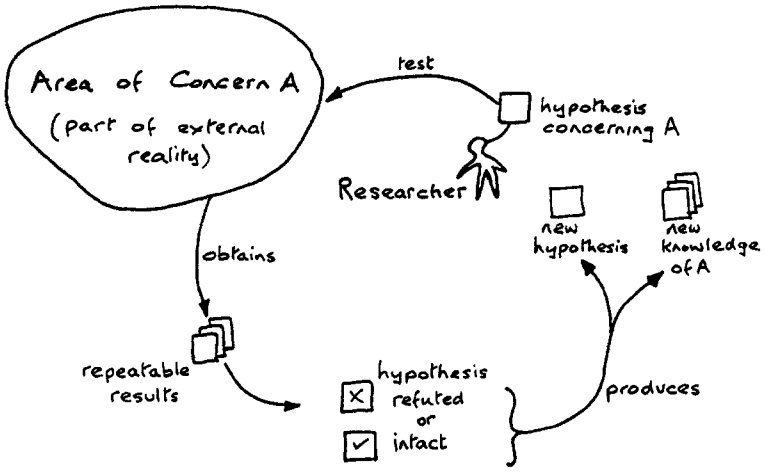

Fig. 1. The hypothesis-testing research process of natural science.

3. THE AR PROCESS AND ITS PROBLEMS

Any research in any mode may be thought of as entailing the elements shown in Fig. 2 (Checkland and Holwell, 1997; after Checkland, 1985, 1991). Particular linked ideas F are used in a methodology M to investigate an area of interest A. Using the methodology may then teach us not only about A but also about the adequacy of F and M. For example, molecular orbital theory and known atomic and molecular dimensions, as a framework F, can be used via a methodology of the computer modeling of the shapes of molecules not yet synthesized in the laboratory to research potentially useful new drugs, an important A. The phlogiston theory of heat (heat as a liquid) is a failed F from the 18th Century. Figure 1 shows the classic form of M for research in natural science.

The change to or modification of F, M, and even A has to be expected in action research. In the 25-year program at Lancaster University which led to the emergence of soft systems methodology (SSM), the initial A was "tackling real-world problems of management" via the application of systems engineering (M). That methodology entailed, as F, *systematic* systems thinking. The methodology was found to be inadequate in that application area, and the learning led to the forming of SSM as a new M, based on a *systemic* F, which was then used for a new attack upon the original A.

This susceptibility to change F, M, and A in research in which the researcher becomes involved in the flux of real-world social situations leads to a (or probably *the*) most important principle in AR. It is a principle almost totally neglected in the literature of this area.

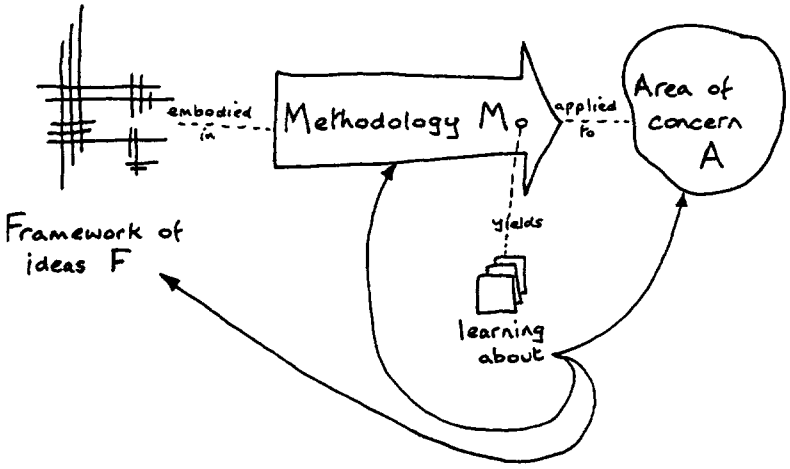

Fig. 2. Elements relevant to any piece of research.

In keeping your intellectual bearings in a changing situation in which the adequacy of F and M and the appropriateness of A are likely to be tested, it is essential to declare in advance the elements F M A in Fig. 2. This is the intellectual structure which will lead to findings and research lessons being recognized as such. Without that declaration, it is difficult to see how the outcome of AR can be more than anecdotal. Many literature accounts of AR leave the reader wondering about the status of that account: How is it to be distinguished from novel writing? To avoid this trap it is essential to define the epistemology in terms of which what will count as knowledge from the research will be expressed. It is the neglect of this principle which leaves AR vulnerable to positivist critics resolutely hanging on to hypothesis testing as a way of researching social phenomena. (They, for their part, have the problem of defining in testable form hypotheses and criteria necessarily expressed in ambiguous abstract terms. The hypothesis that "the designs produced in the Company would be better if design meetings were electronically supported" sounds reasonable if said quickly in everyday language; but every word in it except the articles, the preposition, and the conjunction are richly ambiguous. What would constitute a fair test? Could agreement on the criteria for refutation or survival of the hypothesis be achieved? In social situations one observer's "success" is often another's "failure.")

In constructing another "ideal-type" model of research, this time modified from Fig. 1 to cover AR, we have to accept that the researcher will deal not in hypothesis but in research themes within which lessons can be sought. (In the example above one such theme might be "exploration of design processes" in the Company in question or "support for design meetings.") The researcher interested in particular themes, declaring F and M (from Fig. 2), then enters the "social practice" of a real-world situation in which the themes are relevant and becomes involved as both participant and researcher. It will be necessary to think about that dual role and to negotiate carefully entry into the situation and his or her role in relation to that of participants. Work to effect change and "improvement" (as judged by people in the situation) can then ensue, with the researcher, however his or her role is defined, also committed to continuous reflection on the collaborative involvement and its outcomes. This will entail trying to make sense of the unfolding experience using the declared F and M. This of course may require some rethinking of earlier phases—and again, it is the declared intellectual framework of F and M which allows this to be done coherently. Finally, since real-world situations continuously evolve, the researcher must negotiate an exit from the situation and tease out the serious lessons learnt. This process, as an ideal type, is shown in Fig. 3.

More explicitly, although still as an "ideal type" rather than a prescriptive description, Fig. 3 implies the process model for AR shown in Fig. 4. This

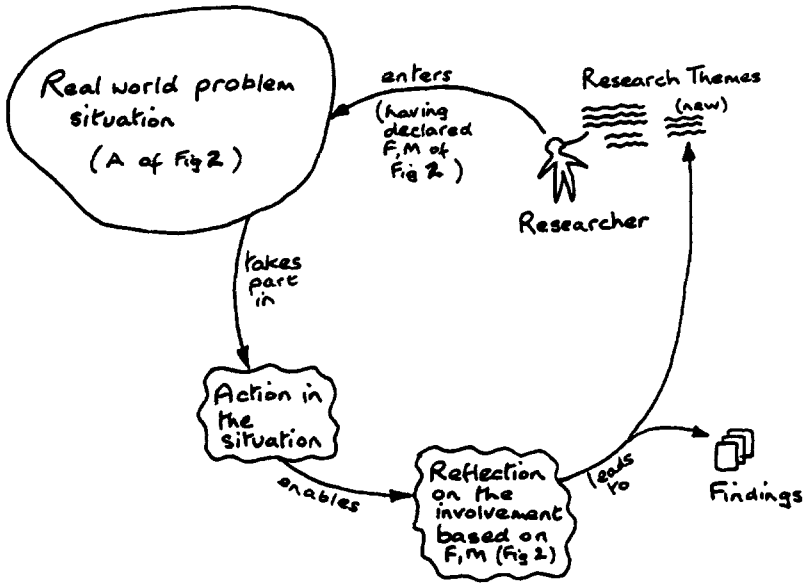

Fig. 3. The cycle of action research in human situations.

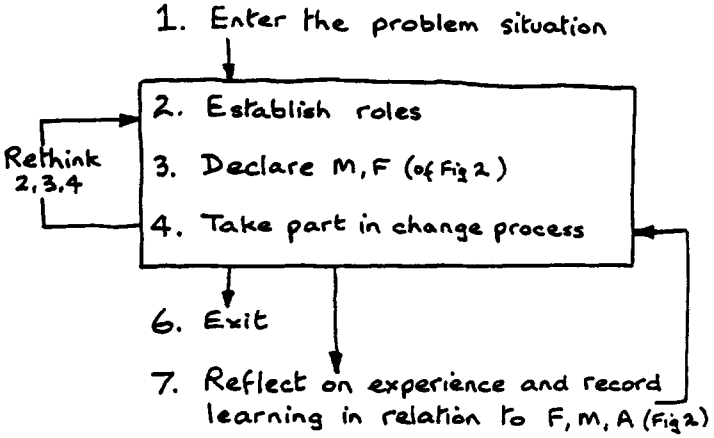

Fig. 4. The process of action research.

covers entering a problem situation, declaring the epistemology in terms of which what counts as learning will be recognized, taking part in the change process, reflecting upon the experience, and recording the learning.

Obviously a process such as that in Fig. 4 could not produce law-like generalizations from involvement in a single situation. In any case AR does not assume that "social laws" await discovery in the same way that physical laws can be regarded as regularities of the universe which do recur whether or not they have yet been noticed and codified. But a serious organized process of AR *can* be made to yield defensible generalizations. For example, a multidisciplinary research team from Lancaster researched the contracting process in the National Health Service in the AR mode and did feel able to generalize. Using the guiding epistemology (F and M in Fig. 2) of soft systems methodology (SSM), a dozen pieces of action research, together with 80 semistructured interviews (structured using SSM-style activity models), were carried out. This did enable defensible generalizations about the gap between the rhetoric of "contracting" and the reality in NHS hospitals and Health Authorities to be made. This work is described by Checkland (1997).

In spite of this kind of evidence that AR can lead to results which can be generalized and transferred to other situations, however, it is obvious that AR cannot aspire to the same claim of validity as that associated with natural science (Campbell, 1988; Phillips, 1992). Achieving credibility, consensus, and coherence does not make a "truth claim" as strong as that derived from replicability of results independent of time, place, and researcher. Action researchers must pay careful attention to the claim of validity relevant to their research into phenomena not "homogenous through time."

It is useful to record the answers implicit in Fig. 4 to hostile questions commonly asked of AR.

3.1. What Exactly Is Being Researched?

The AR process accepts that "themes" have to replace hypotheses. Research in an organization on how to introduce a particular information system, for example, may well evolve into research on what organizational changes are first needed to make it sensible even to contemplate the introduction of a particular system. But themes need to be declared, and a link between them and a putative F and M has to be explicitly argued.

3.2. Who Is Researcher, Who Participant?

The potential merging of the roles "researcher" and "participant" in the situation has to be acknowledged; it should, ideally, be discussed, and the roles may evolve in the course of AR. We have in our practice met among participants

a wide range of degrees of readinesses to take part in the critical reflection on the research as it is enacted, something which is ultimately the researcher's responsibility.

3.3. How Do You Know When to Stop?

The laboratory researcher in natural science can stop when replicable results show that a hypothesis has been refuted or has survived the tests to which it has been subjected. AR as a research mode accepts that social phenomena are "not homogenous through time"; this means that ending a piece of research in an organization is ultimately an arbitrary act. The flux of events and ideas which constitute the research situation will continue to evolve through time. It has to be the researcher's judgment that the chosen methodology (M) and its framework of ideas (F) have yielded significant learning in interaction with the area of application (A). Attempting to "write up" the results will often reveal whether or not a strong case can be made that this position has been reached.

3.4. How Can Results Be Conveyed to Others or Transferred to Other Situations?

Since any organizational situation at a particular time, with its particular participants having their own individual or shared histories, may be unique, it cannot be guaranteed that results can be made richly meaningful to people in other situations. The problem here is not only a problem for AR; it exists also, for example, for those describing case histories. Once again, the importance of the declared epistemology (via F and M) is crucial, though it is neglected as much in case histories in the literature as it is in accounts of AR. For example, Zuboff (1988), in her important and oft-cited work on the profound social effects of computerization in a bank and at a newspaper plant, used participant observation, open-ended interviews, and small group discussions. She wished to understand the interchange between "human responsiveness (feeling, perceiving, behaving) and the experienced 'life-world'" (Zuboff, 1988, p. 243): that was her main theme. She used "an extensive interview protocol designed to translate . . . themes into appropriate questions" (p. 427). Unfortunately her appendix on methodology does not include details of that protocol. With "1500 pages of field notes and transcripts," she "began to build a conceptual map of the territory" (p. 428). Had she included such maps in reporting the work, the reader would have been able to *recover* more of the research content and appraise the judgments being made by the researcher in the course of the work. This criterion of recoverability is sufficiently important to be considered in more detail. If it is met, it will help to justify the generalization and transferability of results from AR (or case study) research.

4. THE IMPORTANCE OF THE "RECOVERABILITY" CRITERION

If we imagine an "ideal-type" spectrum of processes of knowledge acquisition, from experimental natural science at one end to telling stories at the other, then along that spectrum we shall have very different criteria for judging the "truth value" of their outputs or claims. For laboratory experiments in natural science the in-principle "public" repeatability of the experimental happenings, no matter who conducts the experiments, is the basis of the strong criterion which has made natural science the common model of knowledge acquisition. (Ultimately this criterion rests upon the consensus among observers acting in good faith that the happenings are what everyone agrees they are: "This voltmeter is reading 3 volts.") At the other end of the spectrum we shall have the much weaker criterion that this (research) story is "plausible." But that immediately provokes the question, Plausible to whom? What *Weltanschauung* would make this research seem "plausible"? Would another observer find its outcomes "implausible"? The problem for action researchers, knowing that the strong criterion of "repeatability" is beyond their reach, is to do better than simply settle for "plausibility."

Our argument here is that the aim in AR should be to enact a process based on a declared-in-advance methodology (encompassing a particular framework of ideas) in such a way that the process is *recoverable* by anyone interested in subjecting the research to critical scrutiny. The research described by Checkland (1997), carried out by a multidisciplinary team researching purchaser–provider interactions in the National Health Service, attempted to provide such a recoverable process. The work used SSM in the sense-making "Mode 2" form (Checkland and Scholes, 1991), with its particular framework of systemicity. The aim was to make clear to interested observers the thought processes and models which enabled the team to make their interpretations and draw their conclusions. Those observers might not accept the team's interpretations; then a debate about the work could take place which was sufficiently well structured as to be coherent. The weaker "plausibility" criterion does not offer that prospect.

It is the desirability of using the "recoverability" criterion that makes it so important, in the "ideal-type" AR in Figs. 3 and 4, to declare in advance the epistemology in terms of which a piece of AR will acquire what counts as knowledge.

The absence of an insistence on this is the greatest lacuna in the literature of AR, though the point does get made in the more philosophical literature of social science. For example, Phillips (1992, p. 108) argues that if findings are to be taken seriously, they must be supported by appropriate arguments and or evidence; there must be, in the language of Toulmin's account of reasoning (Toulmin *et al.*, 1979), an adequate "warrant" in conjunction with a particular

framework which supplies "backing" for the warrant and, ultimately, the "claims."

> The *claims* involved . . . are . . . *well founded* only if sufficient *grounds* of an appropriate and relevant kind can be offered in their support. These grounds must be connected to the claims by reliable, applicable warrants, which are capable in turn of being justified by appeal to sufficient *backing* of the relevant kind. (Toulmin *et al.*, 1979, p. 27)

The AR literature has rather neglected this kind of consideration. Even as rich an account of AR as that given by Eden and Huxham (1996) does not strongly embrace this point. They usefully set out 15 characteristics of AR, which we drastically summarize below:

(i) researcher intends to change the organization;
(ii) there must be implications beyond the specific situation;
(iii) research seeks theory as an explicit concern;
(iv) any tools, techniques, or models developed need to be linked to the research design;
(v) emergent theory will emerge from both data and initial theory;
(vi) theory building will be incremental and cyclic;
(vii) presentation should acknowledge prescription and description;
(viii) there will be an orderliness in approach;
(ix) exploration of data and theory building should be explainable to others;
(x) later reporting is part of theory exploration and development;
(xi) i–x are necessary but not sufficient for valid AR;
(xii) it is used where other methods are not appropriate;
(xiii) triangulation is used if possible;
(xiv) history and context are given due weight; and
(xv) dissemination of findings goes beyond those involved in a study.

With the exception of point xi (which is a meta-level point, about the set rather than of it), these characteristics map well our experiences in the 25-year action research program at Lancaster. Points v and viii come closest to addressing our concern for the prior declaration of both theory and methodological process, but they need greater emphasis if AR is to deliver more than plausible stories.

5. CONCLUSION

This paper has made an argument which may be summarized as follows.

(1) The repeatability of results in experiments carried out by natural scientists makes the research process of natural science the most easily defended system we have of knowledge acquisition—even though the

repeatability extends only to a consensus on observable happenings, not to any interpretation of them. This success apparently stems from the fact that the natural phenomena studied by natural science are themselves truly regular; they are "homogenous through time" in Keynes' (1938) phrase.

(2) With the increased acceptance that "social reality" is not a given, but is the changing product of a continual intersubjective discourse, there has been in the last decade an increased interest in qualitative research. AR and its many variants are increasingly being treated as serious and appropriate alternatives to the hypothesis testing which is at the core of research methodology in natural science. But proponents of AR need to recognize the limits to the claims they can make for the validity of their approach.

(3) Unable to match the complete replicability of experimental happenings which characterize natural science, researchers investigating social phenomena via AR must at least achieve a situation in which their research process is *recoverable* by interested outsiders. In order to do this it is essential to state the epistemology (the set of ideas and the process in which they are used methodologically) by means of which they will make sense of their research, and so define what counts for them as acquired knowledge.

(4) This gives well-organized AR a "truth claim" less strong than that of laboratory experimentation, but one much stronger than that of mere "plausibility," which is all that much putative AR in the literature can claim.

(5) The literature of AR has so far shown an inadequate appreciation of the need for a declared epistemology and hence a recoverable research process.

Finally, we may remark on this occasion that since systems ideas are a strong component of much methodology which is relevant to qualitative research methods such as action research, it seems appropriate that *Systems Practice* should extend its cover to include interventions aimed at both acquiring knowledge and helping to bring about organizational change.

ACKNOWLEDGMENT

The figures in this paper are reproduced from Checkland and Holwell (1997) by permission of J. Wiley and Sons.

REFERENCES

Argyris, C., Putnam, R., and MacLain-Smith, D. (1982). *Action Science: Concepts, Methods and Skills for Research and Intervention*, Jossey-Bass, San Francisco.

Blum, F. H. (1955). Action research—A scientific approach? *Philos. Sci.* **22**(1), 1–7.

Campbell, D. T. (1988). In Overmann, E. S. (ed.), *Methodology and Epistemology for Social Science: Selected Papers*, University of Chicago Press, Chicago.

Checkland, P. (1981). *Systems Thinking, Systems Practice*, John Wiley, Chichester.

Checkland, P. (1985). From optimizing to learning: A development of systems thinking for the 1990s. *J. Operat. Res. Soc.* **36**(9), 757–767.

Checkland, P. (1991). From framework through experience to learning: The essential nature of action research. In Nissen, H.-E., Klein, H. K., and Hirschheim, R. (eds.), *Information Systems Research*, Elsevier, Amsterdam.

Checkland, P. (1997). Rhetoric and reality in contracting: Research in and on the NHS. In Flynn, R., and Williams, G. (eds.), *Contracting for Health*, Oxford University Press, London.

Checkland, P., and Holwell, S. (1997). *Information, Systems and Information Systems*, John Wiley, Chichester.

Checkland, P., and Scholes, J. (1990). *Soft Systems Methodology in Action*, J. Wiley, Chichester.

Clark, P. A. (1972). *Action Research and Organizational Change*, Harper and Row, London.

Dash, D. P. (1997). Problems of Action Research, Working Paper No. 14, University of Lincolnshire and Humberside.

Eden, C., and Huxham, C. (1996). Action research for the study of organizations. In Clegg, S., Hardy, C., and Nord, W. (eds.), *The Handbook of Organization Studies*, Sage, Beverly Hills, CA.

Engels, P. G. H., and Salomon, M. L. (1997). *Facilitating Innovation for Development*, Royal Tropical Institute, The Netherlands.

Foster, M. (1972). An introduction to the theory and practice of action research in work organizations. *Hum. Relat.* **25**(6), 529–556.

Hult, M., and Lennung, S. (1980). Towards a definition of action research: a note and bibliography. *J. Manage. Stud.* **17**(2), 242–250.

Keynes, J. M. (1938). Discussion of R. F. Harrod's presidential address to the Royal Economic Society. In Moggridge, D. E. (ed.) (1976), *Keynes*, Fontana/Collins, London.

Lewin, K. (1947). Frontiers in group dynamics 11: Channels of group life; Social planning and action research. *Hum. Relat.* **1**, 143–153.

Moggridge, D. E. (1976). *Keynes*, Fontana/Collins, London.

Phillips, D. C. (1992). *The Social Scientist's Bestiary: A Guide to Fabled Threats to, and Defences of, Naturalistic Social Science*, Pergamon Press, Oxford.

Revans, R. (1972). *Hospitals: Communication, Choice and Change*, Tavistock, London.

Susman, G. (1983). Action research. In Morgan, G. (ed.), *Beyond Method*, Sage, Beverly Hills, CA, pp. 95–113.

Susman, G., and Evered, R. D. (1978). An assessment of the scientific merits of action research. *Admin. Sci. Q.* **23**, 582–603.

Torbet, W. R. (1991). *The Power of Balance: Transforming Self, Society and Scientific Inquiry*, Sage, Newbury Park, CA.

Toulmin, S., Rieke, R., and Janik, A. (1979). *An Introduction to Reasoning*, Macmillan, New York.

Whyte, W. F. (ed.) (1991). *Participatory Action Research*, Sage, Newbury Park, CA.

Ziman, J. J. (1968). *Public Knowledge, an Essay Concerning the Social Dimension of Science*, Cambridge University Press, London.