

THE SOCIAL IMPACT OF THE ARTS

Working Paper 1

DEFINING VALUES
Evaluating arts programmes

‘The art of evaluation lies in ensuring that the measurable does not drive out the immeasurable.’

The Audit Commission

François Matarasso

COMEDIA

1996

Defining Values
Evaluating arts programmes

First published 1996 by
Comedia, The Round, Bournes Green, Stroud, Glos GL6 7NL

Publication sales: Jill Chandler 01509 890068

This publication is produced as part of a research programme developed and managed by Comedia. The programme is supported by the Carnegie UK Trust, the Calouste Gulbenkian Foundation, the Esmée Fairbairn Charitable Trust and the Joseph Rowntree Foundation. Other contributing partners include Arts Council of England, Batley Action, Bolton City Challenge, City of Helsinki, Eastern Arts Board, Kirklees Metropolitan Council, London Borough of Hounslow, North West Arts Board, Nottingham City Council, Portsmouth City Council, Scottish Arts Council, Yorkshire and Humberside Arts Board. We extend our thanks to each of these bodies.

ISBN 1 873667 07 8

© 1996 François Matarasso

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher. The right of François Matarasso to be identified as the author of this work is hereby asserted in accordance with the Copyright, Designs and Patents Act 1988

Printed by Performance Print, Nottingham

DEFINING VALUES

FOREWORD	1
INTRODUCTION	1
THE NATURE OF EVALUATION	2
Defining values	2
ETHICAL ISSUES	5
Whose values?	5
ethics and the evaluation process	6
The right to evaluate	9
A CONTEXT FOR ARTS EVALUATION	12
scientific method and Objectivity	13
METHODOLOGICAL ISSUES	15
Quantity and Quality	15
Inputs, Outputs and Outcomes	17
Indicators	18
Cause and effect	19
Replicability	20
PRACTICAL CONSEQUENCES	22
CONCLUSIONS	24
Creativity in evaluation	24
A model for Evaluation	24
SELECT BIBLIOGRAPHY	27

FOREWORD

The Social Impact of Arts Programmes' is Comedia's 4th major study of cultural policy, following research into libraries, parks and the creative city. It addresses key issues in contemporary arts practice, including the social purpose and value of participatory arts, through case studies and related research. The aim of the project is 'to develop a methodology for evaluating the social impact of arts programmes, and to begin to assess that impact in key areas'. This is being addressed by:

- a Establishing a number of case studies to evaluate the social impact of specific programmes, and the assessment structures within which they operate.
- b Reviewing existing literature on social impact in relation to arts programmes, alongside comparable thinking in other fields.
- c Providing a background analysis of the value of arts programmes in achieving social outcomes more commonly targeted through other forms of intervention.
- d Stimulating a debate around the social impact of arts programmes through the publication of working papers, and associated meetings and seminars.
- e Publishing a comprehensive report outlining the findings of the research and proposing a workable methodology for the evaluation of the social impact of arts programmes.

To date, and following a feasibility study in Bolton, the programme includes case studies in Nottingham, Glasgow, Portsmouth, Hounslow, Batley, North Western Scotland, Finland and Australia. A further international case study is looking at the social impact of the creative use of digital technology.

The advisory group members are Ken Bartlett, Franco Bianchini, Tony Bovaird, Roland Humphrey, Alex MacGillivray, Anne Peaker, Usha Prashar, Prof. Ken Robinson, Polly Toynbee, Dr Jill Vincent and Perry Walker. The researchers are Chris Burton, John Chell, Esther Davis, Helen Denniston, Owen Kelly, Naseem Khan, Charles Landry, François Matarasso, Peter Stark and Eva Wojdat.

The study also includes a series of Working Papers, written mostly by people who are not directly involved in the research, but who have specialist knowledge or interest to offer to the debate around the social impact of the arts. As the series title suggests, they often draw on work in progress, or explore issues discursively, without necessarily offering answers.

This Working Paper, no 1 in the series, was written by François Matarasso. It explores the ethical and methodological issues raised by any attempt to assess the social impact of arts programmes. It draws on the theory and practice of evaluation in other areas, especially health care and crime prevention, and concludes with a proposed model for developing arts projects with social objectives.

For further information about the study, please contact François Matarasso on 0115 982 6330 or Charles Landry on 01452 770624

INTRODUCTION

In a world of numbers and quantification, if there are no indicators to assess the value of activities, feelings or relationships these things—however real—have no legitimacy. As Daniel Yankelovich, the renowned American pollster, noted:

‘The first step is to measure whatever can be easily measured. This is okay as far as it goes. The second step is to disregard that which can’t be measured, or give it an arbitrary quantitative value. This is artificial and misleading. The third step is to presume what can’t be measured isn’t really important. This is blindness. The fourth step is to say that what can’t be easily measured really doesn’t exist. This is suicide!’¹

The arts, at least in terms of public policy, have long suffered from the difficulty of quantifying and expressing their value. In the late 1980s an attempt was made by John Myerscough and colleagues to express the financial importance of the arts in a Policy Studies Institute report. This approach fitted well with the preoccupations of the times, and had considerable influence on arts organisations and the public sector generally.

But, despite appearances to the contrary, financial growth is not the only measure of public good. And while the arts certainly can contribute to our GNP, they contribute far more to the health, well-being, stability, development and happiness of British society. The problem is to express these contributions in ways which are clear, provable and helpful in making the most of culture and creative activity.

This paper is part of an international research programme directed by Comedia which seeks to address these problems. Neither this document, nor the study itself, can do more than raise questions, define possibilities and offer some solutions which may be adaptable to particular situations.

Defining Values in particular represents no more than an exploration of the key ethical and practical issues which arise from any attempt to assess the social value of the arts. Because these are clearest in work which has an overt and principal social goal—largely community-based and participatory activity—the paper focuses on this area of activity. However, many of the conclusions expressed are likely to be re-assessed in the light of developing research.

I am very grateful to Dr Jill Vincent of the Centre for Research in Social Policy, Loughborough University; Professor Nick Tilley of Trent Polytechnic; Alex MacGillivray of the New Economics Foundation; Peter Bates of Nottingham Healthcare; Dr Rachel Mason of De Montfort University; and the members of the Advisory Group for much help and advice with sections of this paper. Also to Charles Landry for his usual perspicuity and connections. Naturally, responsibility for the ideas and judgements expressed below remains mine.

François Matarasso

June 1996

THE NATURE OF EVALUATION

Do I know, do you know, what's this thing called man?

God only knows what a man is, I only know his price.

DEFINING VALUES

It is obvious, though not always remembered, that evaluation is fundamentally about values. As this paper will argue, it is not an abstract, quasi-scientific process through which objective truths can be identified. It is necessarily relative. It is always against something. Evaluation without comparison is as meaningless as trying to measure a room without holding one end of the tape measure down.

The important, and essentially political, question about evaluation is which value system is used to provide benchmarks against which work will be measured—in other words, who defines value. This goes far beyond an academic question of assessment. The way in which we define value is part of how we create reality. In economic terms, E. F. Schumacher showed how by not defining coal as a capital asset it was possible to alter not just the balance sheet of the Coal Board, but the global decisions taken as a result.³ In cultural terms, it is argued by many academics and commentators that the very idea of a canon of great art is misleading since it is constructed by a narrow section of the population which effectively defined out the work of other groups, foremost among whom are women. The naming of things creates realities, and those things which are not named, or are mis-named, remain invisible or misconceived.

Policy makers' preoccupation with economic measures, and their reluctance to include social costs and benefits in the balance sheet, is like trying to navigate by compass with a magnet in one's pocket—it draws us off course. We must supplement monetary indicators of value with other more subtle, creative and sensitive benchmarks by which we can measure our progress as a society.

Engaging with evaluation

Many areas of the public sector—from health and education to urban renewal—have accepted the intrinsic value and political necessity of engaging effectively in evaluation. The following justification from the field of social work (James, 1993) can stand for many, in saying the evaluation is important because:

- What gets measured gets done;
- Without measuring results, you can't tell achievement from failure;
- If you can't see achievement, you can't reward it;
- If you can't encourage achievement, you may be encouraging failure;
- If you can't see achievement, you can't learn from it;
- If you can't recognise failure, you can't address it;
- If you can demonstrate achievement, you can win public support.⁴

But, except in the very limited context of funding relationships, the arts world has shown little interest in developing evaluative systems through which to prove its value internally or exter-

nally, preferring, it seems, to state simply that ‘seeing is believing’. Walker (1993), in his history of arts television, offers the customary assertions that the arts world often makes in the absence of actual evidence:

‘In view of the large number of arts programmes that have been produced in Britain since the 1950s ... it is reasonable to assume that a much higher proportion of the British populace is knowledgeable about and interested in the visual arts than was the case in 1950. [...] Television, in all probability, has encouraged many people to visit museums who otherwise would not have done so.’⁵

The vagueness of such claims only confirms the suspicions of those who, seduced by the supposed rigour of their own approach to evaluation, argue that cultural activity is essentially frivolous. The undisguised machismo which divides indicators into ‘hard’ and ‘soft’ simply finds its prejudices confirmed by these arguments.

Assessment beyond aesthetics

But making art is also a way of creating realities, so artists should not be intimidated by engaging with the process of assessment and evaluation. On the contrary, they should feel confident that they have a new way of thinking, and a different set of values which can help us approach policy-making in a more rounded way than before. Creativity recognises and explores the value of subjectivity and the legitimacy of different perspectives. It crosses social and political boundaries, enabling us to respond differently, and to make those leaps of the imagination which are so vital to problem solving.

The rest of this paper looks at some of the ethical and methodological issues this engagement by artists may raise. Other areas of public policy apparently have a longer track record of evaluation (though the arts could claim to have invented the form of evaluation—as subjective as all the others—known as ‘criticism’). It is necessary to understand some of this, and the premises on which it is based, if the arts are to develop their own approach to assessment beyond aesthetics.

CALCULATING WORTH

Although the terms ‘monitoring and evaluation’ have become increasingly common in the arts, at least in negotiation with funding bodies, it is not clear that they are always used in the same way, particularly since they are used by different people with different power relationships.

Evaluation is the process of calculating worth. Its difficulty arises from the essentially relative nature of worth, but it is a process which we carry out continuously. Deciding whether to spend £30 on a sweater, a couple of CDs, a meal or shares in British Gas is a personal evaluation, in which we calculate the relative worth, to us, of the benefits to be purchased.

Evaluation is not, despite being widely used in this way, shorthand for ‘How did we do?’ though answering that question is a step along the way.

Determining the effectiveness of work undertaken is an internal management process. By using monitoring systems, well-managed organisations can have a clear understanding of their performance, set against their own objectives and standards. This is not the same as determining its wider value. The production of one theatre company may be of more value than

that of another, perhaps more competent group: the impact (or value) of the work may be only partly connected with the quality of the management, or even the craft skills of the performers, director and other staff.

The company's objectives and standards will have been established in response to the expressed or implied demands of people outside the organisation—customers, government, shareholders and others. Arts organisations set their standards, consciously or not, to reflect the demands and interests of customers (paying or not), funders, their peers in the arts world, and their own staff. The relative weight given to each of these 'stakeholders' varies widely, though a degree of balance between them is evident in most successful organisations. Most of the inevitable and natural argument about the work of arts bodies reflects a struggle between the competing, or even conflicting, interests of the stakeholders.

Evaluation is crucial to every arts body, not because a particular funder requires it, but because it is the process through which all important policy and planning decisions are made. And, as the rest of this paper makes clear, it is fundamentally about values.

EVALUATION STAGES

The evaluation process is cyclical but includes a number of distinct stages, integrated with the conventional stages of a planning cycle. These are: setting objectives; planning how to meet them; undertaking the work; and reviewing progress. There are more complicated matrices for expressing this, but the basic principle is clear: you decide what you want to do and how to do it, before carrying out the work and assessing your performance against your original objectives.

But what looks attractively simple in a management handbook becomes much more complicated when applied to the real world. And as far as arts organisations are concerned the process often becomes bogged down because it is so difficult to review progress without an agreed framework for doing so. The problem is that the work of arts organisations is ultimately about quality and assessing it requires the definition of quality.

Defining quality depends entirely on your values. What one person sees as being high-quality may strike another, with a different view of the world, as being a waste of money. The relative positions in the process—and the power relationships—of the stakeholders is liable to influence or even to determine their assessment of a piece of work. So at this stage, the apparently objective process of planning and evaluating the work, slips away like an eel.

As a result, most formal evaluation reports of arts work become narratives which describe what happened, but rarely relate back to the objectives or offer evidence that these have been met. The rest of this paper explores some of the ethical and methodological problems which arise from this.

ETHICAL ISSUES

The most special part of the human being is the ability to create values. Values don't fall out of the sky. They aren't given to us. We don't find them in nature. We create them ourselves.

WHOSE VALUES?

Since evaluation is fundamentally about value, any engagement with it raises the question as to whose values are being adopted. It was noted above that there are stakeholders in every arts project or organisation. In most cases these will be the customers, the funders and the staff, but there may be other groups more peripherally concerned including politicians, the media and pressure groups. In some cases these are able to have a major, even terminal impact: a project to celebrate the opening of Wansbeck District Hospital in Northumberland with local choirs was abandoned after a press campaign and political objections based on a misconception of the source of the money. Clearly, the values of different stakeholders can have fundamental influence on the way an initiative is seen, or even whether it goes ahead.

The values of the various stakeholders influence the work of arts organisations in different ways. In the case of funders there is an ongoing dialogue, a formal application procedure, and a process of negotiation with which both sides are familiar. The interests of staff are represented through management procedures, while those of other arts professionals are filtered in through informal channels. Finally, paying customers will choose whether or not to buy the experience offered by an arts organisation. In the end, if the values of a theatre company are too much at odds with those of their potential audiences, the problem will become evident and force some type of solution.

IN THE ABSENCE OF MARKET VALUES

If things are complicated in conventional types of arts provision like theatre, orchestras and galleries, it becomes much more difficult where the 'customer' does not pay and may not even choose to participate. Community-based projects frequently do not charge admission, while people may find themselves caught up in an event simply because they live in a particular area, attend a day centre or use a hospital. And it is these projects which most commonly have an expressed intention to change those with whom they have contact. Negotiations between community arts projects and their funders are commonly based on claims about the social impact of the work, and how participants will become more confident, skilled and independent or even less criminal.

There is a fundamental ethical problem here: is it right to seek to produce change in another person without their informed consent? For someone with liberal democratic values, the answer must be no. It may be argued that informed consent can only be secured through a developmental process, and that consent can be given at each stage. But even if consent to change were negotiable, it remains the case that, with very few exceptions, the participants and intended beneficiaries of arts projects are not involved in setting the objectives, or in determining the criteria against which success will be measured. They are, in effect excluded

from participating in determining the process of change, although it may have lasting impacts on them as individuals or communities.

THE POLITICAL AGENDA

Evaluation has become a central concept in British politics in the past decade, as the nature of the contract between the public sector and the citizen has changed. By introducing market forces, or quasi-markets, the political right has sought to renew the individual's power through a form of consumer rights. The most recent and obvious example has been the creation of the Citizen's Charter by the Government since 1991. This has begun to establish a series of national benchmarks applying to public or publicly-funded services, against which the individual consumer can judge what he or she gets. But the opposition has argued that the granting of consumer rights has not been matched by similar political rights, and the transformation of citizens into customers effectively avoids the issue of who makes decisions in the first place.

If, to the right, the ideas behind charterism are to expose services to greater competition through the disciplines of the marketplace, to the left the need is to enhance and justify public services, together with much wider access to information. Both share the tools of quality auditing, customer contracts, and redress. The concern for quality in services leaves unanswered the issue of power; that is, who has the power to take decisions, to assess standards of service and to evaluate outcomes. (Hill, 1994)⁷

It must be recognised, therefore that evaluation and impact assessment, independent of the philosophy of social research, has an important political context.

ETHICS AND THE EVALUATION PROCESS

Evaluation has only two organisational purposes—to improve the organisation's ability to meet its objectives effectively, efficiently and economically, and to demonstrate the organisation's value to its stakeholders. These are aims which any organisation in receipt of public funds can be expected to endorse. That many are suspicious, even hostile to evaluation, underlines the fact that its application is far from neutral.

The report of the Law Centres Federation London Evaluation Scheme (Thornton, 1992) makes clear that evaluation undertaken internally, albeit with an external facilitator, is helpful and welcome, while evaluation by funders, perceived to be driven by unexpressed political objectives, and over whose terms the subjects have no say, is met with suspicion, and may not deliver any lasting benefit.⁸ Given the essentially social and political context for evaluation in the public sector, ethical considerations are of great importance, as has been recognised in recent years by the development of codes of practice for researchers in the social sciences and other disciplines.⁹

The key issue is the different outcomes which evaluation may produce for the subject and the commissioner. This has been expressed as follows:

The costs/benefits ratio is a fundamental concept expressing the primary ethical dilemma in social research. In planning their proposed research, social scientists have to consider the likely social benefits of their endeavours against the personal costs to the individuals

taking part. Possible benefits accruing from the research may take the form of crucial findings leading to significant advances in theoretical and applied knowledge. [...] The costs to participants may include affronts to dignity, embarrassment, loss of trust in social relations, loss of autonomy and self determination, and lowered self-esteem. [...] The process of balancing benefits against possible costs is chiefly a subjective one and not at all easy. (Frankfort-Nachmias and Nachmias, 1992)¹⁰

While social scientists may proceed carefully, weighing the potential benefits against costs, in practice the evaluation required by funders of arts programmes does not offer the opportunity to say no. A researcher may argue that

Being ethical limits the choices we can make in pursuit of truth. Ethics say that while truth is good, respect for human dignity is better, even if, in the extreme case, the respect of human nature leaves one ignorant of human nature.¹¹

But failing to evaluate a Council funded project on ethical grounds will not cut much ice in the committee chamber. So the successful and ethical evaluation of arts projects *must* take proper account of the rights of all the individuals involved.

INDIVIDUAL PERFORMANCE

It is also essential to distinguish between the performance of individuals and organisations, and the impact produced. Creating a community play in Cirencester may be much easier than doing it in Moss Side. There are useful parallels with the recent move towards teacher assessment and school league tables. In particular it has been argued with some force that simply ranking schools by exam results gained is neither fair nor helpful, since the location, intake and resourcing of different schools varies widely. Higher GCSE pass rates in a suburban selective school than in a city-centre comprehensive may have nothing to do with the schools' actual performance, since they are not starting from the same point. In other words, what matters is performance in context.

But evaluating performance in context can only be part of the solution, because it is an organisational approach. The rights of the individual worker, whether artist or teacher, need to be considered. In the United States, where there is a longer experience of teacher assessment, a number of ethical principles have been developed. One model which might be adapted to the context of arts workers in social contexts was originally set out by Strike (1990).¹² It is based on a series of seven principles:

- 1 **Due process:** Procedures and standards must be reasonable, known and applied consistently.
- 2 **Privacy:** Information about an individual should be confidential, should be used in ways known to the person, and should be limited to job-related aspects of their activities.
- 3 **Equality:** Decisions should not be made on irrelevant grounds (e.g. gender, ethnicity etc.)
- 4 **Public perspicuity:** The evaluation processes should be open to the public.
- 5 **Humaneness:** Consideration is due to the feelings of those involved in evaluation processes.

- 6 **Client benefit:** The interests of the ‘client’—in this case, students, parents and the public—must come before those of organisations and their staff.
- 7 **Academic freedom:** Evaluation must not be conducted in a way which inhibits the basic purpose and values of education.
- 8 **Respect for autonomy:** Individuals are entitled to exercise reasonable professional judgement in regard to their work.

These principles could easily be adapted to reflect the situation of artists and arts workers engaged in projects intended to produce a social impact. Academic freedom should obviously be replaced by artistic freedom or integrity, but otherwise few changes are needed. More significant perhaps, are the implications for procedural and managerial change in developing equitable standards in evaluating the work of individuals in the overall task of impact assessment.

PARTICIPANTS IN EVALUATION

Safeguarding the rights and interests of arts workers in an evaluation process is important although, where the evaluation of a publicly-funded arts programme is concerned, a willingness on the part of the workers to subject their work to external evaluation comes with the territory. But that evaluation necessarily involves the participants and intended beneficiaries of the programme. The extent to which the evaluation may acceptably impose demands or even costs on them, and the nature of those costs, is a much more complex issue.

The concept of informed consent (and its counterpart, informed refusal) has been developed to deal with some of these ethical issues. (It is interesting to note, in passing, that artists have no scruples in not offering informed consent to people who may figure in work from photographs to the use of intimate letters and even suicide notes.)¹³ It is a mark of good practice in some areas of social research, to provide participants with written conditions and guarantees governing their involvement, but this does not happen in the evaluation of arts programmes.

The conduct of social research has also been the subject of much thought, both in the United States and the UK, and is now circumscribed by legislation (e.g. the Data Protection Act 1984) and by such administrative mechanisms as ethical review committees and codes of conduct agreed by professional bodies. Questions of privacy, anonymity, confidentiality, and the legitimacy of deception have been closely analysed and there is a substantial body of published material.¹⁴

Sensitivity in handling the outcome of evaluation is also critical. The damage which may be done to participants in a project subsequently deemed to have failed may outweigh any benefit resulting from changes. While this is not an argument for allowing poor quality work to continue, the longer-term effects of the presentation of findings must be considered. The impact of being publicly branded a failing school was devastating on a primary school in County Durham.

THE RIGHT TO EVALUATE

A number of fundamental ethical questions arise from this discussion, which centre on the right of different stakeholders to contribute to the evaluation process on their own terms. This does not mean being consulted or asked for an opinion within a framework established by others, but the opportunity to contribute to the shaping of that framework in the first place. It means ensuring that different values and viewpoints are accommodated within the process. It may be that the values of different stakeholders are irreconcilable, but if this is the case, making it clear can only help avoid latent conflict.

SELF-ASSESSMENT AND EMPOWERMENT

Empowerment is commonly stated as the ultimate aim of arts programmes with social objectives. It does not seem possible to develop the empowerment of individuals or communities without giving them control over their own process of change. Some professionals will question the role legitimacy and value of involving participants in self-determined evaluation. But there is a precedent in the field of education as outlined by Ross in *Assessing Achievement in the Arts*, following a programme of research and case studies. Exploring the place of discussions about their work with pupils, he writes:

‘What came as a real surprise, was not only the readiness of the pupils to seize the opportunity to talk about their work presented by the conversation format, but also their considerable, latent capacity for exploring, explaining and evaluating the aesthetic experience through talk. Assessment in the arts ceases therefore to begin and end with the teacher’s perceptions, judgements and statements and turns right around to restore the work to its maker, the pupil. And so a new role emerges for the teacher: to equip the students with the reflective skills to monitor and assess their own work.’¹⁵

The implication is that evaluation is itself a part of the creative, artistic process (a point made slightly differently by Eisner¹⁶) and should be legitimised as such. This, after all, is the type of questioning process which professional artists engage in continuously.

ACCESSIBLE CONCEPTS AND LANGUAGE

Involving people in assessment and evaluation presents another challenge to professionals, since their concepts and language must be made accessible to all those involved. The measurement of health status provides a clear example of how different things can look from the point of view of the professional and the patient. Seale (1993) illustrates how differently disability can be described by doctors and disabled people. A purely physical approach is used in the Barthel Index in Maryland, where evaluative tools have been developed by professionals to assess levels of disability:

The Barthel Index

Feeding; Personal toilet (wash face, comb hair, shave, clean teeth); Ascend and descend stairs; Controlling bowels.¹⁷

But many disabled people, whatever their view of disability politics, are more concerned with the consequences of their disability on their social well-being—on how they experience disability. The Nottingham Health Profile (Hunt et al., 1986) was developed by lay people to measure how they felt, and it concentrates on how illness or disability affects them. In each

case, the patient answers yes or no to the questions, and a numerical value is assigned to the response. The indicators are strikingly different:

Nottingham Health Profile I'm tired all the time; I find it hard to bend; Everything is an effort; I feel lonely.¹⁸

Clearly the conceptions of need, benefit and how to measure held by professionals and patients are fundamentally different, as is the language used: where the Barthel Index is exclusive and dehumanising, the Nottingham Health Profile is open and inclusive. But the needs of both patient and health worker have a legitimacy which has to be recognised through the care models they develop. The answer here, as elsewhere, can only be to integrate the various perspectives of different stakeholders in an activity. The problems arise again, however, if we consider what weight each stakeholder's views should be given: should the ratio between a patient's feelings and a doctor's ideas be 50/50, or 70/30, or 20/80? The answer depends on which position we occupy.

Whatever the issues of health care, it is clear that the arts could certainly benefit from exploring issues of legitimacy of viewpoint and, equally importantly, the principle of setting relevant and understandable indicators.

PEER REVIEW

Medical audit was originally imposed on doctors in the USA by insurance companies seeking to reduce their costs. It is defined by the Department of Health as 'the systematic, critical analysis of the quality of medical care, including the procedures used for diagnosis and treatment, the use of resources, and the resulting outcome and quality of life for the patient'.¹⁹ Medical audit can involve a number of methods, including peer review, where the work of one group of doctors is assessed by another. This in many ways mirrors practice in the arts, where assessors used by the Arts Councils and Regional Arts Boards tend to be artists or arts professionals.

While many doctors and artists would argue that this approach is better able to understand and evaluate the work, its professional focus may be questioned. Peer review is at risk of concentrating on process rather than outcome, because it is the process which interests those involved, and they have sympathy with the professional difficulties experienced by colleagues. In some areas of French theatre assessment practice, the quality of the production has sometimes seemed less important than the ideas of the *metteur-en-scène*.

In the last 3 or 4 years, therefore, a shift towards multi-disciplinary clinical audit has taken place, and some leading edge services are involving service users in designing and delivering audits. This echoes cultural evaluation work undertaken in Australia. In assessing community arts projects, Deidre Williams used an observer group which brought together people who could be expected to have different perceptions of the project. They included:

- 5 people in positions of community leadership or business;
- 5 people who were only involved in the artistic work;
- 5 people who were audience members or who only saw the finished product;
- 5 people who are family members of the participants;
- 5 people who supported the project in some way but did not get involved.²⁰

This approach certainly seems worth exploring, though the identification of the members of such a group would naturally vary with the project and the purposes of the evaluation.

STAGES OF PARTICIPATION

In conclusion, it is worth noting the work of Andrea Cornwell on different approaches to participation practice.²¹ The relationships of evaluator and participant can usefully be explored in this context.

MODE OF PARTICIPATION	TYPE OF PARTICIPATION
CO-OPTED	Tokenism; manipulation; representatives are chosen, but have no real input or power
CO-OPERATING	Tasks are assigned, with incentives; but outsiders decide agenda and direct process
CONSULTED	Local opinions are asked for; outsiders analyse and decide on a course of action
COLLABORATING	Local people work with outsiders to determine priorities; responsibility, however, remains with outsiders for directing the process
CO-LEARNING	Local people and outsiders share their knowledge to create new understanding and work together to form action plans, with outsider facilitation
COLLECTIVE ACTION	Local people set their own agenda and mobilise to carry it out, using outsiders not as initiators/facilitators but as required by the local people

A CONTEXT FOR ARTS EVALUATION

Intelligence, ‘brains’, in the political mind is still almost exclusively associated with verbal and mathematical reasoning and with propositional knowledge. Practical work in the arts is not so much associated with intelligence, which schools are required to develop, as with talent, which they might choose to encourage.²²

There is a widespread assumption (in Britain at least) that the arts and the sciences are different, even opposed systems of belief and ways of acting. This contributes to a lack of confidence which people working in creative industries can feel in the face of apparently more ‘scientific’ disciplines. Fully recognising the subjective nature of much of their work, artists and arts workers often see other disciplines and professions as characterised by a degree of objectivity to which it would be futile even to aspire. This is certainly evident in the feelings of many arts professionals about evaluation, but it is striking to find it reflected even in an area like education:

‘The particular value of scientific research in education is that it will enable educators to develop the kind of sound knowledge base that characterises other professions and disciplines; and one that will ensure education a maturity and sense of progression it at present lacks.’ (Cohen & Manion, 1994)²³

To an outsider, this is an oddly un-self-confident view to find expressed in a 400 page volume describing research methods in education, and which boasts a bibliography running to almost twenty pages. If this reflects a field lacking in maturity and a sense of progression, how might research into the social impact of the arts be characterised?

THE EVALUATION OF HEALTH CARE

In fact the position of evaluation in other professions and areas of public policy is not as impressive as it may seem from the world of the arts or even education. Health care is perhaps the field from which one could expect best practice in terms of scientific assessment guiding practice. Medicine is, after all, a science. It is also a matter of life and death, where there should be no room for approximation. Everyone knows the principle of drug trials in which only part of the group receive a new medicine. But a brief review of the evaluation of health care reveals some striking inconsistencies:

- Evaluation is required for new drugs, but not for new techniques, many of which—especially in the area of psychiatric medicine, have never been formally evaluated.
- Some treatments have never been evaluated because they are known to be effective, and it is therefore considered unethical to deny them to some patients for the purposes of conducting a comparative trial.
- Clinicians do not have an unblemished record of responding to evaluation: for many years surgeons continued to treat breast cancer by mastectomy, effectively ignoring research evidence that less radical surgery (lumpectomy) was just as effective.²⁴

- If the evaluation of interventions is uneven, policy assessment is far behind it. One of the key elements of the BMA's campaign against the 1991 NHS reforms was the argument that the internal market had not been evaluated.²⁵

The reality is that the quality of health care evaluation varies as widely as its influence on policy and practice in the NHS. As the medical sociologist Clive Seale has admitted,

It may be unrealistic at the policy level to hope for the direct linkage of results to decisions in the way that the rational, step-by-step approach suggests. The results of research may often influence policy makers in a more diffuse way, seeping into their thinking, in a manner that eventually provides enlightenment.²⁶

Other areas of public policy do not have better records than medicine. It is not just the performance of the economy which is the subject of debate, but, as the New Economics Foundation and others have argued, the indicators set up to measure that performance which are flawed.²⁷ As Simon Kuznets, mastermind of the GNP, admitted, 'the welfare of a nation can scarcely be inferred from a measurement of national income [such as GNP].'²⁸ In these and other fields as diverse as social work, crime prevention and environmental protection, the processes of evaluation are as controversial as they are important.

The rapid pace of change in public policy merely underlines how rarely it is based upon a careful weighing up of scientific evidence. One might question, as the BSE disaster develops, whether the question of scientific evidence is merely a useful excuse to avoid taking action.²⁹ The problem of course, is not that scientific evidence does not exist (though that may be the case), but that the very processes by which it is produced result in it being limited in scope and hedged about with caveats. Its meaning and use to guide action are so contested that its value as an aid to development of public policy is severely restricted.³⁰

If scientific research has its weaknesses, it may be argued, it is still the best tool we have. Certainly the arts can learn a lot from the scientific approach, and especially the experiences of the social sciences, but they equally have much to offer in developing sensitive, creative, people-centred approaches to evaluation which begin to address the outcomes, rather than the outputs, of policy initiatives.

SCIENTIFIC METHOD AND OBJECTIVITY

Despite the weaknesses in existing evaluation, and the way those weaknesses are commonly ignored by policy makers, it is still widely felt that scientific methods are invariably superior to all others. Science cornered the market in objectivity a couple of centuries ago, and it will take more than this paper to redress the balance. The *beau ideal* of scientific evaluation is the randomised controlled trial (RCT), originally developed as a means of testing agricultural technology in promoting crop growth, and boosted by an epidemiologist named Archie Cochrane in his book *Effectiveness and Efficiency* (1971). The concept, of randomly dividing patients into two groups (to distribute their characteristics randomly) and giving them different treatments (which are not known to them), is well known in its application to the testing of new drugs, though ethical questions have already been referred to.

PROGRAMMES AND POLICIES

But it is significant that randomised controlled trials have not been widely used to test policies or programmes, despite Cochrane's suggestion that it be used to evaluate social work. There are problems in seeking to extend RCT's from the precisely-defined scope of drug assessment to public policy:

- 1 It is very difficult to control the conditions and variables in the programme, and therefore to establish cause and effect.
- 2 Political problems exist in securing the co-operation of a control group for a programme, since they must agree to take part, but not to receive whatever benefit the programme is perceived to have.
- 3 It is not possible to replicate the 'placebo effect' in relation to programmes since they know whether or not they are participating.

Cochrane recognised the weaknesses of the RCT in evaluating diseases where improvement or deterioration had to be measured subjectively (as is the case in psychiatry) and lamented that 'we have so far failed to develop any satisfactory way of measuring quality'.³¹

OBJECTIVITY

The question of objectivity, that touchstone of scientific authority, further confuses the issue. Philosophers and social scientists have written extensively on the issue of objectivity in social research, creating conflicting schools of thought which have in their turn been subjected to acute criticism.³² But it is not necessary to engage with these abstruse arguments to recognise that the day-to-day business of programme evaluation is fundamentally subjective. Social policy generally, and arts projects specifically, are value-driven. This is not the abstract world of theoretical physics where, in principle at least, the search for truth outweighs personal considerations.³³ The evaluation of individual programmes can have life-altering impacts on everyone involved, from those who use them or gain their salary through them, to officers or politicians who have endorsed them. Everyone has a vested interest and objectivity in such a climate is likely to be elusive.

But this does not devalue the process. Instead it requires evaluation procedures to take account of the legitimate subjectivity of different stakeholders. By recognising that the same activity or programme can be perceived differently, depending on one's point of view, it may be possible to create a composite picture which is, if not the truth, at least a reliable basis for further action.

SCIENTIFIC PROCESS AND THE ARTS

Given that it has not proved possible to apply scientific methods of evaluation to all health care, let alone other areas of public policy, and in recognition that findings are not always used as a guide to action, the evaluation of arts programmes is not as weak as it might at first appear. It would certainly seem that there is more to learn from research and evaluation models which recognise the limitations of scientific method. Of these, the closest to the type of evaluation already used by arts projects are 'before-and-after' models, such as those used by urban regeneration programmes like City Challenge. Self-evidently, however, these depend on the availability of 'before' evidence, something which is unusual in the arts.

METHODOLOGICAL ISSUES

The art of evaluation lies in ensuring that the measurable does not drive out the immeasurable.

The Audit Commission*

QUANTITY AND QUALITY

THE ATTRACTION OF FIGURES

All evaluative studies rely on two types of information: quantitative evidence, expressing the size of various measures, and qualitative evidence, which expresses more obviously subjective judgements about worth. The increasing emphasis on evaluation in the public sector has relied heavily on quantitative evidence (albeit of increasing complexity) because it is easier to get, easier to use, and appears authoritative. Take, for example, the following statements

- People aged over 65 account for 48% of NHS expenditure.
- Those over 85 (1% of the population) account for 12% of expenditure.

They are typical of the type of quantitative evidence which appeals to politicians. Their primary characteristic is their portability: lodged in a corner of the mind, they can be pulled out at any appropriate moment to back up a point in discussion with the appearance of authoritative finality. (In fact, the second is wrong: people over 85 account for 8% of expenditure, but who could challenge the figure given?) The Policy Studies Institute study of the *Economic Importance of the Arts* was an effective advocacy document because it produced memorable, apparently authoritative ‘facts’ of which people could make immediate use.³⁵

Bodies advocating certain courses of action often use polling organisations to provide them with ammunition in the form of quantitative information. If it were reported that 84% of the population supported the National Lottery, the drama of the figure would obscure questions about alternatives, or how the question was phrased.³⁶

PERFORMANCE INDICATORS

This approach to quantitative evidence can be seen in the performance indicators used by City Challenge and Single Regeneration Budget (SRB) companies to report to the Department of the Environment. Bolton City Challenge Baseline Study gives a clear picture of the economic, environmental and social problems they were established to tackle—the ‘before’ evidence against which progress will be assessed. Its statistics range from size of business by turnover and residential property type, to reported crime and infant mortality rates. From this information a series of ‘impact indicators’ have been established. So for Strategic Objective 7 (‘Reduce crime and improve the social fabric’), the following indicators are set:

- Decrease in reported crime; (Baseline: 3,380)
- Decrease in residential burglary; (Baseline: 694)
- Decrease in theft of or from cars; (Baseline: 1,201)
- Decrease in survey respondents’ fear of home break in; (Baseline: 78%)³⁷

Being able to compare the level of these and other key indicators after 5 years is a real contribution to determining the impact of the programme.

The Community Forest Programme uses similar indicators. The annual monitoring data includes more than 60 separate measures, but they are restricted to easily measurable, usually physical, change. For example 27 of the indicators specify the number of hectares affected by the programme. The next largest group of indicators relates to financial and other resources secured for the programme. Only four indicators relate to community involvement:

- Events targeted at farmers;
- Volunteer days on practical projects;
- Community events (publicised events, organised by all the partners for the local community e.g. festival, arts event, guided walk). Excludes school events.
- Number of schools involved in the forest (includes Christmas card recycling)—where school has been in contact with any of the partners.³⁸

THE WEAKNESSES OF QUANTITATIVE INFORMATION

While these indicators are certainly useful, they can only tell a very partial story. Perhaps most significantly, they support and perpetuate a reliance on capital (rather than revenue) programmes to produce change evident in many areas of British public policy.

They also rely on what is easily measured. So environmental indicators used by Bolton City Challenge touch only on physical matters (e.g. the number of pedestrian crossings) rather than local people. Other environmental indicators might have included percentage of residents who walk or cycle to work; are members of conservation groups; use local parks and green spaces; are involved in local nature conservation projects; and much else.

But the critical issue is the difficulty of addressing questions of quality through this approach. A body might meet its target of delivering 10,000 training weeks, without producing any lasting outcomes, because the training on offer was inappropriate or bad.

Quantitative information only pretends to objectivity. Its weakness lies both in the unspoken or unrecognised assumptions which underlie it and in its susceptibility to manipulation, recognised as long ago as Disraeli ('Lies, damned lies and statistics').

APPROACHES TO QUALITY

It is not surprising, therefore, to find that efforts have been made to balance statistical information with other more subtle forms of qualitative measure. Gathered through a range of research methods, including focus group discussions, interview, participant observation, action research etc., this material is much harder to organise, evaluate and present in useful form. But its importance should not be underestimated. Market research—a 'hard' discipline if ever there was one—makes extensive and sophisticated use of this type of evidence.

Quality measures have been introduced by the health service, because simple cost effectiveness statistics can be misleading. It is normal to compare the costs of a treatment or programme with the benefits obtained. The result is expressed in terms of a cost/effectiveness ratio: for example £200 per person cured of a duodenal ulcer by drug therapy, compared to

£300 per person cured of a duodenal ulcer by surgery.³⁹ Cost effectiveness is relative, setting the achievement of identical benefits against each other and showing that one can be achieved more cheaply than another. Problems arise where unlike benefits are compared: e.g. the benefit of not dying (derived from heart transplants) with the benefit of reduction in disability (derived from hip replacements). The health sector has therefore developed the concept of 'life years' to even out these differences by providing a standard of measurement. This allows comparison of different forms of treatment in terms of cost per life-year saved.

But prolonging life with increased disability may not be the most desirable result. So a refinement is the 'quality-adjusted life year' (QALY) designed to improve on a crude concept of life at any cost. For example if a treatment programme delays death by six months but leaves the patient in pain or bedridden, then that 0.5 of an additional life year can be 'weighted', so that in units of QALYs it amounts to less than six months. Similarly, if a treatment improves physical status or reduces the duration of pain or disability, the change in weights will result in an increased number of QALYs.⁴⁰ But this approach has not been universally adopted because:

- Different people have different views on the valuation of health states.
- No account is taken of external factors, such as dependants.
- The arithmetic suggests equivalences which may not exist.
- Some doubt the wisdom of making such decisions clear, explicit and apparently rational.

QUALITY AND THE ARTS

Could this approach be applied to the arts? Certainly the cost-effectiveness ratio has a relevance to arts programmes with social objectives. It would, for example, be possible to compare the cost-effectiveness of crime prevention work with young people which used the arts and sport, provided indicators of the outcomes were available. But trying to adjust these basic comparisons for quality seems over-complex and unnecessary. If these methodological problems exist in the scientific world of medicine, they appear almost insuperable in the arts, where the subjective, personal response is central to the meaning. It is not merely acceptable for the same work to elicit a range of responses from passionate enthusiasm, through indifference to hostility, it is intrinsic to its nature. There is no universally admired art, and that which most closely approaches it is likely to be at least a century or two old. In the end, gathering evidence of quality in an arts programme may require similar processes to market research: a combination of discussion, and recorded assessments from the spectrum of stakeholders.

INPUTS, OUTPUTS AND OUTCOMES

The emphasis on performance indicators in the public and independent sectors has tended to concentrate on inputs and to a lesser extent, outputs. As a result community-based arts groups are becoming used to setting objectives and targets and evaluating their progress. To date, this has largely been an internal process. Having set the objective for example, of running 40 open workshops with an average attendance of 8 people in a 3 month period, groups are assessing

and reporting on that basis. This is a useful management tool, but it is restricted to inputs and outputs: i.e. what resources were applied and what was produced as a result. It scarcely touches on quality and still less on outcome, which is to say, what happened as a result of the outputs being achieved. What effect did these workshops have on the participants?

Some groups have begun to move in this direction. First Take, for example, who offer community video training in the North West, record the outcome of their courses in terms of the participants' future career development—i.e. whether they find jobs, do further training, return to unemployment and so on. This is a good start, but it is evident that assessing outcomes less clear-cut than employment status requires further thought. A more detailed discussion of the relationship between inputs, outputs and outcomes will be found in the companion working paper, 'Creative Accounting' by the New Economics Foundation.

THE ABSENCE OF OUTCOMES

But the first difficulty which many arts projects with social objectives face is a lack of clarity about what outcomes are actually intended. Although most of those involved have a more or less clear sense of why they are doing what they are doing, it is rarely formalised. It is therefore easy to lose sight of it and make the wrong decisions, or to find the work being pulled in different directions by unexpressed purposes. This lack of clarity is common in organisations of all types. For instance, the overall aim of local authorities is to enhance the quality of life of people who live in their area. Stimulating the local economy to create employment is a means to that end, but, since it is that which people are engaged with on a daily basis, it is easy for it to become confused with the aim. Such confusion can lead to controversial or inappropriate decisions as employment is pursued even to the point of damaging overall quality of life. If arts initiatives with aims beyond purely aesthetic concerns are to assess and demonstrate their achievements in those areas, they will have to be much clearer about the outcomes which are intended.

INDICATORS

Ensuring that outcomes are clearly stated is only a first step. To determine progress towards those outcomes, some sort of constant scale to measure against, is required. Comedia has been developing ways of using indicators to measure outcomes, and made preliminary suggestions in a discussion document published for the Arts Council in 1993.⁴¹ The use of indicators is not in itself new, as the discussion of City Challenge performance underlined. But this also underlined that their use has related primarily to inputs and outputs, rather than outcomes which, as we have seen, are the real concern of an organisation.

Indicators need therefore to be rethought so that they can relate more closely to what a project's stakeholders really want to achieve. So, if a community festival was intended to have an impact on social cohesion in an urban neighbourhood, one might use indicators such as:

- Changes in the patterns of social networks of participants and non-participants.
- Growth of membership of and participation in community groups, and of informal and formal partnership between them.

- Changes in patterns of criminal behaviour among participants (e.g. recidivism) and in the local area (including incidents of harassment etc.).
- A reduction in time spent by participants alone and at home.
- Developing understanding of different cultures among project participants.
- Changes in local media reporting of the area.
- Changes in relationship between local authorities and constituents.

UNFORESEEN AND NEGATIVE IMPACTS

Indicators would need to be set as a project is planned to have real validity. However, given the unpredictable nature of cultural activity, it is essential to avoid pre-determined indicators preventing the identification of unforeseen benefits or impacts. The problem here is that there will not usually be a benchmark figure against which unforeseen impacts can be measured.

A further complication is the question of negative impacts. (This is not the same as the sort of negative results produced by a poorly planned or executed piece of work, since they will appear in relation to the indicators.) The methodological problems arise where the objective of a programme is to achieve the absence of something—e.g. a reduction in cases of racial harassment. The problem is a bit like proving that the dog you bought to deter burglars is responsible for the fact you haven't been burgled, or whether it is simply good luck.

CAUSE AND EFFECT

The next stage in the chain of methodological difficulty is that being able to show change in relation to a pre-defined indicator does not prove that the change was produced by the programme being evaluated. The immediate answer to this problem is to use a control group, although, as touched on above, control groups for social programmes present real difficulties. So the normal route of evaluation seeks to establish a causal link between the programme and its outcomes by the elimination of outside factors. This approach encourages the evaluator to remove as many variables as possible, in a search for *internal* validity, i.e. the clearest link of cause and effect.

In practice, absolute internal validity is unachievable in the context of an arts programme. But, even where it can be shown that a given intervention produced a given result, this approach to evaluation cannot answer the question '*Why* did it succeed or fail?'

CAUSAL MECHANISMS

Pawson and Tilley developed a model for assessment of crime prevention programmes which explores their impact on causal mechanisms in a given context, rather than simply seeking links between an intervention and an outcome (Tilley, 1993).⁴² They argue that, since the particularities of each situation affect the outcome of any programme, they must be taken into account rather than ruled out of the valuation process. For example, the introduction of security cameras into a shoppers' car park, might have a discernible impact on the number of offences committed there. It would be easy to assume that the introduction of security cameras is *per se* a useful crime prevention measure, although in practice their use in a commuters' car

park may have no effect. The placing of cameras in the first car park may make shoppers feel more secure, and therefore more inclined to use the car park, resulting in an increase in hourly activity through the day. This may make potential offenders feel that the risk of being interrupted by someone is becoming too great, and therefore discourage them from taking that risk as often as before; (though their activity may be displaced elsewhere, rather than suspended). But in the commuters' car park the cameras will have no impact on activity during the day time, so the benefits will not be replicated. In such circumstances, a better intervention might be to encourage a car washing company to work in the car park as a means of deterring offenders.

It is argued therefore that only analysis of the causal mechanisms triggered by a given programme can show why it has succeeded or failed. However, evaluation processes usually bypass these complexities, trying only to establish a causal link between the programme and its outputs.

REPLICABILITY

More critically, an evaluation which pursues internal validity by disregarding local conditions cannot predict the replicability of a particular programme or intervention. However close it gets to internal validity, it cannot achieve external validity. Although the RCT may be an appropriate model for isolating the precise effect of a treatment on a specific condition, it cannot cope with the infinite variability of the body of society.

People seeking to replicate earlier, successful programmes evaluated through causal link analysis rather than causal mechanism analysis are often disappointed to find that they cannot get the same results. Although the variations arise naturally from the different contexts in which the work is being repeated, they can lead people to reject a valid programme or approach.

This has been clearly shown by Nick Tilley in his review of attempts to replicate the successful Kirkholt Burglary Prevention Project. In considering why other programmes often failed to reproduce this success, Tilley isolates a number of key situational characteristics which did not apply in follow-up projects. He suggests that, since subsequent projects cannot offer strict replication of the original Kirkholt project, it is possible for defenders of Kirkholt to explain differences in outcome by arguing that adherence to the model was not close enough. He reaches a number of useful conclusions in respect of the conduct and reporting of evaluations in all social programmes:

- 1 Where model projects are envisaged the conjectures relating to context, measures, mechanisms and outcome patterns should be explicit.
- 2 People considering replicating interventions, should consider the context into which they are to be made.
- 3 Evaluators should ensure that they compare like with like, and consider 'what brought it about'.
- 4 Replication evaluation should be of paramount interest to social scientists.⁴³

These considerations will be obvious to arts workers who have been tried to reproduce a successful art project. The project contains a huge number of variables:

- The personalities, skills, experience and qualities of the artists involved.
- The aims, agendas and relationships of the project stakeholders.
- The location and social situation of the work and the resourcing available.
- The personal and social characteristics of the participants.
- The timetable, planning, even the time of year.

Differences, which are inevitable, in any of these will alter the nature and eventual outcome of an arts project. Therefore, drawing on the Pawson-Tilley model, one should be looking for the key elements of the project which are essential to its successful replication. For example, how much were the personalities involved a factor in the success? Could a different arts worker achieve the same results? How would changing the location or frequency of the project, or the type of art work affect future outcomes? Once again, as in the case of causal links, the only practical solution is to make an informed judgement. In this context it may be the way the arts were explored, rather than the medium used, or the way in which working relationships between those involved were developed.

Creative initiatives cannot have internal validity. It is never going to be possible to produce a model arts project which can be applied to social problems like a tablet. But that is their strength—it is in the creative unpredictability of their outcomes that arts projects add an essential tool to the range of social action.

PRACTICAL CONSEQUENCES

There has been a cultural change in the voluntary sector in the past 15 years, driven partly by growing contractual demand for its services, and partly by an increasing requirement by funders for evidence of its effectiveness in addressing that need. Securing grant aid from local authorities, government departments and charitable trusts has become more competitive, and the need to evaluate services has grown. Most voluntary organisations have begun to undertake some sort of internal evaluation, often at the insistence of funding agencies.

Among them, arts organisations, and especially community-based arts bodies, have been increasingly required to evaluate their work. A full consideration of the practical consequences arising from the increasing need to evaluate the social impact of arts projects is inappropriate here, but a few salient points can be noted.

RESOURCES

The voluntary sector is, by its nature, running up the down escalator. There is so much to do, with such inadequate resources, that it is hard to see evaluation as a priority except, and paradoxically especially, when it is a requirement of the funder. Methodical evaluation is increasingly required of larger, Government funded programmes like City Challenge, or the Safer Cities programme, (though the use to which the results are put may be questionable), but is not yet generally required of arts projects. Unless and until evaluation and impact assessment are seen as a central part of the practice of arts projects with social objectives, they will not receive either the resources or the attention required

EXPERIENCE

In exceptional cases, evaluation is undertaken in partnership with an academic institution or independent consultants, but the great majority of evaluation is internal, and therefore amounts to a formalised reporting back on work undertaken. But artists and arts workers, who are so often responsible for evaluating their work, generally do not have experience nor training in the relevant skills. This is entirely understandable—they are employed as arts workers, after all—but it remains true that if more and better impact assessment is to be undertaken, those responsible for doing it will need full professional support.

DISSEMINATION

Finally, there is the problem of dissemination: what to do with the outcomes of evaluation. In practice, even using them to make internal change within the organisation they apply to is difficult. Making them available more widely for others to benefit from is often impossible.

The arts, which are taught on an individual, skill-focused basis, do not have the type of academic structures to support such understanding. The various professional magazines—*Mailout*, *Sounding Board*, *Animated* etc.—go some way to filling the gap, but have a limited readership, and limited experience in this field. The publication of reports without effective means of distribution (or even peer review) is an unreliable way of advancing debate and knowledge. Finally, the time and interest of the potential readership is also highly variable.

A good example of these weaknesses combining in practice is the Safer Cities programme where several hundred projects produced scarcely any valid or useful results, despite the explicit requirement of evaluation being built in from the start.⁴⁴

If the social potential of arts programmes is to be better understood and more widely exploited as an instrument of public policy, serious consideration will have to be given not only to the production of reliable and usable evidence, but to how that material is to be disseminated within and, perhaps more importantly, beyond the world of the arts.

CONCLUSIONS

CREATIVITY IN EVALUATION

Even this partial review of the ethical and methodological issues which surround evaluation suggests a number of conclusions:

- 1 Full engagement in evaluation processes is critical to arts bodies, both in their own interest and to contribute to the shaping of values in the public domain.
- 2 Although other areas often have more experience, the arts have different, more creative approaches to offer.
- 3 Over-zealous pursuit of scientific objectivity, and the internal validity of evaluation processes, is inappropriate and unhelpful approaches to the evaluation of social programmes and especially arts projects.
- 4 New approaches to evaluation which reflect the co-operative and creative values of the arts can and need to be developed.

In many ways these conclusions echo the thinking of Raymond Illsley, who acknowledges the difficulties inherent in a scientific study even in the field of health care. He argues that 'successive partial evaluations and reforms are superior to perfect trials which demand such stringent conditions that they cannot be carried out.' His 'illuminative approach' distinguishes between processes and outcomes to analyse the causal chain in health-care activity. If the processes which lead to outcomes are clearly defined, then it is possible to set standards for their delivery. In turn the attainment of these standards may in many cases be measured fairly simply.

A MODEL FOR EVALUATION

Drawing on this, and on the practical experience of a feasibility study undertaken by Comedia for North West Arts between November 1995 and February 1996, I offer some preliminary conclusions about how arts projects intending to produce and demonstrate social impacts might be developed. These conclusions are summarised below, and will inform some of the further case studies in the research project. They draw on the following principles:

- Projects intended to produce social benefits should address stated needs or aspirations.
- It is unethical to seek to produce change without the informed consent of those involved.
- The needs and aspirations of individuals or communities are best identified by them, often in partnership with others, such as local authorities, public agencies and arts bodies.
- Partnership requires the agreement of common objectives and commitments (though not all goals need be shared by all partners).
- Those who have identified a goal are best placed to ascertain when it has been met.
- An arts project may not be the most appropriate means of achieving a given goal.

A FIVE STAGE PROCESS

Bearing these in mind, the following process is suggested as an appropriate approach to developing projects and determining their impact, (though other ways of working may be equally valid).

1 Planning

The first stage is to establish a partnership between the stakeholders and identify the problem, need or aspiration to be addressed. Having done this, it is possible to devise an appropriate arts project or decide that some other form of intervention would be more appropriate. The partners should include potential participants, the organisation or department developing the work and, depending on circumstances, others agencies (e.g. funders). Arts workers will often become involved only after it has been decided that an arts project is needed.

2 Indicators

Having set objectives relating to identified needs, and a strategy to meet them, the partners can decide how they will know when they have been met. The thinking already done will make it much easier to set indicators linked to project goals, and plan how to collect the information needed. Allowance must also be made to accommodate unforeseen outcomes.

3 Execution

By this stage, everyone involved in the project will know what is being done, why, and how it is to be assessed, and have given their consent to these. The work can be carried out, and monitored by whatever method has been chosen as most appropriate. It may be that the stakeholders wish to amend the project aims, though this should be through a rider to the stated aims, since it is tempting to change the objectives, or the way they are interpreted, in the light of the way the project is proceeding. It is preferable to use any discrepancy between original and subsequent aims to explore the expectations of those involved.

4 Assessment

On completion of the project, sufficient evidence should be available through the monitoring process to compare with the chosen indicators, to enable everyone to assess progress. Time will have to be set aside to collate it and to consider other, unforeseen, impacts.

5 Reporting

Shortly after completion, the different stakeholders can set out their assessments of progress and discuss common ground or discrepancies through a joint discussion. Reporting is essential both to close the project, and to discuss and begin planning potential routes for further joint or individual development. The final stage ought thus to become the first stage of the next cycle of the process.

Developing a project in this way should not only help to produce evidence but contribute positively to the way in which people work together, since the first two stages will require that the aims and commitments of all those involved are stated clearly and generally agreed.

EMPOWERMENT THROUGH SELF-MANAGEMENT

It is not suggested that this process is easy, although it is certainly easier than sorting out the negative impacts of a mis-handled arts project. The ability of individuals and community

groups to identify and prioritise needs, and to devise strategies to address them, will vary and may need to be developed. The skills needed to facilitate such a process will also need to be developed.

But this is the very empowerment that so many community-based arts projects are intended to deliver. It only requires a shift in perception so that the whole process, from first contact to discussing what to do next, is understood to be the arts project.

The five stages set out above are deliberately cyclical, returning the partners not to the point where they began—since they have all developed and learnt since then—but to the starting point of the next stage, where the process can be repeated, more effectively, by people who gone through it once. It should lead to increasingly independent and autonomous activity, since it gives people control over their own participation and development.

SELECT BIBLIOGRAPHY

The following list includes publications which have been used in preparing this Working Paper, together with some which have not been seen but which may be of interest.

- Allen, D. (1981) 'An Analysis of the factors affecting the development of the 1962 Hospital Plan for England and Wales' *Social Policy and Administration*, 15(1) pp-318.
- Belson, W.A. (1986) *Validity in Social Research*, Gower Publishing, Aldershot
- Burgess, R. (ed.) (1989) *The Ethics of Educational Research*, Falmer Press, Lewes
- Cohen, L. & Manion, L. (1994) *Research Methods in Education*, Routledge, London.
- Crane, J. A. (1982) *The Evaluation of Social Policies*, Kluwer-Nijhoff, Lancaster.
- D.E.S. (1995) *Art in the National Curriculum HMSO*, London.
- Davey, B. & Popay, J. (eds.) (1993) *Dilemmas in Health Care*, Open University Press, Milton Keynes.
- Davey, B., Gray, A. & Seale, C. (eds.) (1994) *Health and Disease: A Reader*, Open University Press, Milton Keynes.
- Eisner, E. (1979) *The Educational Imagination: On the design and evaluation of school programmes*, New York.
- Feek, W. (1988) *Working Effectively—A Guide to Evaluation Techniques*. Bedford Square Press, London
- Feek, W. and Smith, D.J. (1983) *Value Judgements*, National Youth Bureau, Leicester
- Frankfort-Nachmias, C. & Nachmias, D., (1992) *Research Methods in the Social Sciences*, Edward Arnold, London
- Goldberg, E.M. & Connelly, N. (1982) *Effectiveness of Social Care for the Elderly, An Overview of Current and Evaluative Research*, Heinemann/PSI, London.
- Goldberg, G. et al (1983) *Monitoring Voluntary Action*, Association of Researchers in Voluntary Action and Community Involvement, Wivenhoe.
- H.M.I. (1990) *The Teaching and Learning of Drama; An Inspection Report*, HMSO, London.
- Hersen, M. A. and Bellack, A. S. (1988) *Dictionary of Behavioural Assessment Techniques*, Pergamon, New York.
- Illsley, R. (1980) *Professional or Public Health?* Nuffield Provincial Hospitals Trust, London.
- Kingsley, S. and Taylor, M. (1983) *Action-Research: Myth or Reality?*, Community Projects Foundation, London.
- L.C.F. (1988) *Questions of Value: a Framework for the Evaluation of Law Centres*, Law Centres Federation, London
- Lishman, J. & Smith, G. (eds.) (1984) *Pluralistic Evaluation*. *Evaluation Highlights No 8* Dept. of Social Work, University of Aberdeen pp 140-162.
- Mason, R. (1993) *Artists in Education*, De Montfort University, Leicester.
- May, T. (1993) *Social Research: Issues, Methods and Process*, OUP, Milton Keynes
- O'Hear, P. & White, J. (eds.) (1993) *Assessing the National Curriculum*, Paul Chapman, London.
- Parker, L. (1995) *Music Against Drugs, an evaluation of a drugs prevention diversionary activity*, HMSO, London
- Patton, M.Q. (1982) *Practical Evaluation*, Sage (London)
- Popper, K. (1959) *The Logic of Scientific Discovery*, Hutchinson, London.
- Richards, H. (1985) *The Evaluation of Cultural Action*, Macmillan, London.
- Ross, B. & Channon, G. (1985) *Evaluating Your Organisation*, Research Training Initiatives, Newcastle.
- Ruddock, R. (1981) *Evaluation: A Consideration of Principles and Methods*, Manchester Monographs, Manchester.
- Shirtcliff, C. (1994) *Proving Your Worth: a framework for evaluating telephone advice, tribunal representation and group work*, Law Centres Federation London
- Thornton, C. (1992) *London Evaluation Scheme: Final Report*, Law Centres Federation London
- Tilley, N. (1993) *After Kirkholt—Theory, Method and Results of Replication Evaluations*, Police Research Group, Home Office, London
- Walker, J. A. (1993) *Arts TV: A history of arts television in Britain*, John Libbey, London

References

- 1 As quoted in *The Quality of Life of Cities*, Comedia 1995
- 2 Bertolt Brecht, **Supply & Demand**, translated by Eric Bentley.
- 3 E. F. Schumacher, (1973) **Small is Beautiful**, Blond and Briggs, London.
- 4 'Inside Quality Assurance', Ann James, *Community Care* 27 May 1993
- 5 Walker, J. (1993) *Arts TV*, p.194.
- 6 Joop Doorman, Professor of Philosophy, Erasmus University, Rotterdam, *Dasarts Dialogues 5*, Stichting Dasarts, Amsterdam 1996
- 7 Hill, D. (1994) *Citizens and Cities*, Urban Policy in the 1990s, Harvester Wheatsheaf, Hemel Hempstead. p.6.
- 8 Thornton, C. (1992) *London Evaluation Scheme: Final Report*, Law Centres Federation London
- 9 Cohen, L. & Manion, L. (1994) *Research Methods in Education*, Routledge, London. p. 347.
- 10 Frankfort-Nachmias and Nachmias (1992), in Cohen, L. & Manion, L. (1994) *ibid.* p. 350.
- 11 Cavan (1977), in Cohen, L. & Manion, L. (1994) *ibid.* p. 359
- 12 Strike, K. (1990), 'The Ethics of educational evaluation' in J. Millman and L Darling-Hammond (eds.), *A New Handbook of Teacher Evaluation* (Sage, Newbury Park, California) in Cohen, L. & Manion, L. (1994) *ibid.* p. 377.
- 13 There are two conflicting beliefs here, which are paralleled in science and the arts. Each, in Western philosophy, abrogates to itself the right of free and uninhibited enquiry in search of self-determined truth. Problems arise when this right of free enquiry collides with other, equally firmly held beliefs: for example, the sacred nature of Australian Aboriginal art, or, to remain in the cradle of Western liberalism, a belief in the inalienable right of human dignity.
- 14 Burgess, R. (ed.) (1989) *The Ethics of Educational Research*, Falmer Press, Lewes
- 15 Ross, M. (1993) *Assessing Achievement in the Arts* p.156 ff.
- 16 Eisner, E. (1979) *The Educational Imagination: On the design and evaluation of school programmes*, McMillan, New York, p. 187.
- 17 Clive Seale, 'The Evaluation of Health Care' in *Dilemmas in Health Care*, ed. Basiro Davey and Jennie Popay, Open University Press 1993.
- 18 Hunt, S. M., McEwen, J. and McKenna, S.P. (1986), *Measuring Health Status*, Croom Helm, London; in Seale, *ibid.*
- 19 Department of Health, *Medical Audit*, Working Paper no 6. HMSO London 1989
- 20 Williams, D. (1995) *Creating Social Capital*, Community Arts Network of South Australia, Adelaide.

- 21 Towards Participatory Practice: PRA and the participatory process, Draft Paper, Andrea Cornwell, Department of Anthropology, SOAS, London 1995; in *Participation in Poverty Reduction*, Gert Lüdeking, United Nations Centre for Human Settlement, Volume II No 1, 1996.
- 22 Robinson, K. *The Arts in the National Curriculum: Which Way Now?*
- 23 Cohen, L. & Manion, L. (1994) *ibid.* p. 40.
- 24 Seale, *ibid.*
- 25 Alastair Gray, 'Rationing and Choice' in *Dilemmas in Health Care*, ed. Basiro Davey and Jennie Popay, Open University Press 1993.
- 26 Seale, *ibid.*
- 27 See *Creative Accounting*, Sanjiv Lingayah, Alex MacGillivray and Peter Raynard, Comedia 1996.
- 28 *If the GDP is Up, Why is America Down?*, Cobb, C., Halstead, T. & Rowe, J., *Atlantic Monthly*, October 1995, pp.59-78.C ited in Lingayah et al. *ibid.*
- 29 In May 1996, as Government explained it was waiting for scientific evidence before taking action, the Ministry of Agriculture revealed that it was spending £1 million a week less on scientific research than in 1985 when BSE was first reported. (*Independent* 22 May 1996).
- 30 I am grateful to Dr Jill Vincent of Loughborough University for help with this section.
- 31 A.L. Cochrane, *Efficiency and Effectiveness; Random Reflections on Health Service*, in Davey B. et al, *Health and Disease*, Open University, Milton Keynes.
- 32 See for example, *Social Research, Issues, Methods and Process*, Tim May, OUP 1993.
- 33 Peter Bates drew my attention to problems with objectivity even in theoretical physics, as expressed in the Heisenberg Principle of Uncertainty. The problem is that, in order to look at a sub-atomic particle, it is necessary to shine a light on it. But that light bashes the particle and moves it, so the very act of observation distorts the original position.
- 34 Quoted in Thornton, C. (1992) *London Evaluation Scheme: Final Report*, Law Centres Federation, London.
- 35 Gordon Hughes argues in a critique of Myerscough's work: 'Thus, if a case is to be made for the economic importance of the arts it must rest upon the non-market benefits of artistic activities rather than conventional measures of the scale of the sector's employment and output'. He goes on to calculate that the more narrowly defined arts sector accounts for less than 0.2% of GDP: (*Policy Studies* Vol. 9. no.3, 1989).
- 36 There is a parallel with the use of images to convey messages: they can also be very powerful, and apparently possess authority, but are entirely open to manipulation. Even arts organisations have been known to select the best photographs to tell their story.
- 37 Bolton City Challenge Baseline Study, Final Report, (Bolton Metro, March 1995); p. 55
- 38 Community Forest Annual Monitoring Data, Community Forest Unit.
- 39 Seale, *ibid.*

40 Scale, *ibid.*

41 *The Social Impact of the Arts, A Discussion Document*, Comedia, Stroud, 1993. See also **The Creative City, Working Paper3, Indicators of a Creative City**, Franco Biamnchini and Charles Landry. Comedia 1994.

42 Ray Pawson, Leeds University and Profesor Nick Tilley, Nottingham Trent University: See Tilley, N. *After Kirkholt—Theory, Method and Results of Replication Evaluations*, Police Research Group, Home Office, London 1993

43 Tilley, N. *After Kirkholt—Theory, Method and Results of Replication Evaluations*, Police Research Group, Home Office, London 1993

44 Information from Professor Tilley.

WORKING PAPERS AND ASSOCIATED TITLES

Comedia is pleased to announce the publication of the first 3 working papers in the Social Impact series. The first two deal principally with the theory and practice of evaluation, drawing on experience in social and economic policy. The third reports on the experience of development agencies in using cultural programmes. Further working papers are in preparation and will be published in the autumn.

All working papers cost £5, plus 50p postage: The first 3 are available for £15, post free.

1 DEFINING VALUES: Questions of evaluation in the arts

François Matarasso

30 pp.

Defining Values explores ethical and methodological issues arising from trying to assess the impact of arts programmes. It looks at the question of values, and the critical importance of viewpoint, making a case for the participation of intended beneficiaries not only in evaluation, but in determining its nature. The paper addresses methodological questions including the difference between outputs and outcomes, causal links, indicators and replicability, drawing on the practice of evaluation in health care, crime prevention and other areas. It concludes with a proposed model for developing arts projects.

François Matarasso is a Comedia principal and director of the social impact study.

2 CREATIVE ACCOUNTING: *Beyond the bottom line*

Sanjiv Lingayah, Alex McGillivray & Pete Raynard

40 pp.

The paper offers a critique of conventional economic indicators from the national to the organisational level, and looks at how cost/benefit analysis has been used at project level. The authors explore new approaches to expressing changes in quality of life at each level, and look at how these might be used in arts assessment, before concluding with some suggestions about assessment.

Sanjiv Lingayah, Alex McGillivray and Pete Raynard work in the indicators and social auditing programmes of the New Economics Foundation.

3 THE ART OF SURVIVAL: *Investigating creativity in humanitarian aid and development*

Helen Gould

20 pp.

The paper gives an overview of a little-known field—the use of the arts as a tool of social change in the developing world. Based on the author's own extensive researches, *The Art of Survival* draws together an inspiring series of examples from Africa, South America, the Indian sub-continent and elsewhere.

Helen Gould is a journalist: staff writer for The Stage Helen and has written for various arts publications and national newspapers. She is raising funds for the 1st international conference on Arts, Human Rights, Aid and Development in London in November 1996 in alliance with the International Movement of Rights and Humanity

THE ART OF REGENERATION: Urban Renewal Through Cultural Activity

Charles Landry, Lesley Greene, François Matarasso and Franco Bianchini

80 pp.

Cultural projects have played an increasingly important role in British urban regeneration since the mid-1980s, but recent developments have focused less on prestige capital projects, and more on the capacity of arts activity to support community-led renewal. This study describes more than 100 initiatives, including 14 detailed case studies from the UK and abroad.

Price £15.00