

RESEARCH ARTICLE

Theory-based evaluation and the social impact of the arts

Susan Galloway*

Centre for Cultural Policy Research, University of Glasgow, Glasgow, UK

The well-documented challenges in researching the social impacts of the arts are closely related to key issues in contemporary social research and evaluation, most particularly the problem of causal attribution. The article contends that some of the most common criticisms of the evidence base for the social impact of the arts relate to the successionist model of change which underpins positivist social science research and evaluation. Illustrating this with reference to research on the arts and quality of life, the article considers the alternative generative understanding of causation that underpins theory-based evaluation (TBE) approaches, favoured recently in the UK as part of the “What Works?” agenda. While these approaches fit well with knowledge about the determinants of arts impact, the article considers whether in fact TBE approaches offer an effective strategy for understanding how and why arts engagement can result in social change. The limitations and possibilities of TBE are considered with reference to four recent UK studies of the impact of the arts on individuals which make use of them.

Keywords: arts; social impact; theory-based evaluation; realistic evaluation; Theory of Change; quality of life

Introduction

Belfiore and Bennett (2006) have reconnected the contemporary policy debate about the social impact of the arts to its historical philosophical roots. In a similar vein, the widely acknowledged methodological challenges involved in researching the social impacts of the arts have largely been discussed in detachment from wider debates about evaluation and social science research methods that cross public policy areas and disciplines. Many reviews have concluded that there is little or no evidence to support the claims made about arts impact, but without acknowledging how widespread an issue “lack of evidence” is in contemporary public policy and the history of discussion about this within ongoing debates about “what works?” (Davies, Nutley, & Smith, 2000; Sanderson, 2000).

The empirical evidence for the social impact of the arts has been reviewed and appraised by numerous authors (including the Australian Expert Group in Industry Studies – AEGIS, 2004; Belfiore & Bennett, 2007; Burns Owen Partnership, 2005; Coalter, 2001; Guetzkow, 2002; Oakley, 2004; Reeves, 2002; Ruiz, 2004; Shaw, 1999; White & Rentschler, 2006). This includes reviews of the evidence of arts impact drawing on literatures in areas such as health (Staricoff, 2004), education (Winner & Hetland, 2000), criminal justice (Hughes, 2004), mental health (White & Angus, 2003), social inclusion (Jermyn, 2001; Newman, Curtis, & Stephens, 2003) and regeneration (Evans & Shaw, 2004; Landry, Bianchini, Henry, & Brookes, 2004). Identified

*Email: s.galloway@arts.gla.ac.uk

deficiencies in the evidence base have been summarized, and summaries of the summaries produced (including Belfiore, 2006; Reeves, 2002; Ruiz, 2004).¹

Sanderson has identified a “tendency to emphasize methodology rather than theory as the basis of ‘good evaluation’”, ascribing it to what he calls the “dominant rationalist-modernist paradigm” (2000, p. 439). Certainly critical discussion of social impact studies has tended to concentrate on technical rather than ontological or epistemological issues (for example AEGIS, 2004, p. 28; Cowling, 2004, p. 131; Reeves, 2002, p. 103). Most of this work concludes by urging investment in larger scale, longitudinal research and evaluation, and by calling for more “robust, innovative and transparent methodologies” to be developed (White & Rentchler, 2006, p. 12).

Arguably the main issue for advancing our understanding of the effects of arts interventions is ontological; it is not research methods but the most effective “orientation” or “logic of enquiry” (Pawson, Greenhalgh, Harvey, & Walshe, 2004, p. 2). The largely unanswered question is what types of research *approach* are best suited to investigating the social effects of the arts? This article addresses this issue by appraising the usefulness of theory-based evaluation (TBE) as one possible way forward.

Definitions

The understanding of social impact used here is broadly that set out in the 1993 discussion document by Landry, Bianchini, Maguire, and Worpole (1993). This defines impact as “a dynamic concept which pre-supposes a relationship of cause and effect” (p. 8), adding that the study of impact in an arts context “is essentially about the transformative power of the arts in terms of personal and social development . . . along a continuum from totally negative to totally positive” (p. 9). They offer the following working definition of the social impact of the arts:

Those effects that go beyond the artefacts and the enactment of the event or performance itself and have a continuing influence upon, and directly touch, people’s lives. (p. 50)

In the UK, discussion of the social impacts of the arts has become closely associated with the social inclusion agenda of the New Labour government with its particular focus on participative arts activity. The politics of this has been critiqued elsewhere and is not the focus of discussion here (Belfiore, 2006; Merli, 2002).

Methodology

A literature search was undertaken of academic and other specialist databases, the websites of arts and cultural organizations and networks as well as contacts with experts to locate articles in which TBE had been used in the context of arts impact. Locating evaluation work using a theory-based approach proved difficult because of the range of terminology in use and the fact that these are seldom adopted as keywords. Other terms in use include “theory-driven” approaches, programme theory, programme logic, causal models, theory-of-action and intervention logic. Interviews were also undertaken with research practitioners with direct practical experience of applying TBE within this context.

Article structure

The article has four main sections. Part 1 considers how weaknesses in the evidence base for the social impacts of the arts reflect common problems in public policy evaluation and social science research. The key question of causal attribution is discussed in Part 2, which considers the view that many of the problems identified with the “evidence base” for arts impact are actually to do

with the dominant successionist model of causation. Part 3 considers the alternative generative model of causation that underpins realist social research and TBE approaches, such as realistic evaluation and Theory of Change. The final part discusses the potential of TBE approaches for researching the social impact of the arts with reference to four recent UK arts studies which adopt this approach.

Part 1: Critiques of the evidence base

The now substantial literature appraising the “evidence base” for the social impact of the arts has identified numerous weaknesses (AEGIS, 2004; Belfiore, 2006). These weaknesses, to do with environmental and structural constraints, evaluation capacity and quality, causal attribution and complexity, although much discussed in relation to arts research, do not represent a particular “failing” of research into the arts (Oakley, 2004). Rather the requirement for “evidence” of the effectiveness of policy interventions has raised a common set of issues across public policy.

This is borne out by the wider evaluation literature, which ascribes some of the common weaknesses in the evidence base for social interventions to structural problems, to the policy environment, the type of public management, and the nature of the policy cycle (Barnes et al., 2005; Evans, 2005; May, 2006; Pawson, 2006, p. 8). As in the arts, political imperatives influence not only the setting of targets and outcomes, but can often intervene in or influence the research or evaluation process itself (Judge & Mackenzie, 2002, pp. 300–312; Shipman, 1997). These factors also create difficulties for evaluation in areas such as health, where research budgets are substantially higher than the arts (Blackman et al., 2006; Judge & Mackenzie, 2002).

Failure to identify the longer term impacts, rather than outputs, of arts activities, including unforeseen or negative impacts, are key criticisms made of the evidence base for social impact of the arts (Coalter, 2001; Belfiore, 2002; Shaw, 1999). But the difficulties in actually achieving this are a commonly identified issue in the measurement of impact across a range of public policy areas using traditional data-driven evaluation approaches (Burton, Goodlad, & Croft, 2006; Mackenzie & Benzeval, 2005; Oakley, 2004).

Evaluation capacity and methods in the area of arts impact are generally considered to be underdeveloped (AEGIS, 2004; Belfiore, 2006; Coalter, 2001; Oakley, 2004). Again, in this respect the arts are not exceptional; recent reviews of the evidence base in relation to public health (Millward, Kelly, & Nutbeam, 2003) and area-based regeneration (Thomson, Atkinson, Petticrew, & Kearns, 2006) amongst others have reached similar conclusions about these areas of public policy intervention.² And, as in the arts, these reviews have often produced fairly critical assessments of research quality, examples being criminal justice (Harper & Chitty, 2005) and education (Hillage, Pearson, Anderson, & Tamkin, 1998; Tooley & Darby, 1998), where a vigorous debate has ensued (Brown, 2005).

But the central criticism made of social impact studies in the arts, and perhaps the main reason for the “no evidence” conclusion, is the failure to “prove” causation between arts participation and particular effects (AEGIS, 2004; Belfiore, 2006; Merli, 2002; Reeves, 2002; Winner & Hetland, 2000). Explaining and attributing change and the questions of paradigm and methods this presents is a fundamental concern of the disciplines of social science and evaluation and the much documented methodological problems inherent in researching the social impact of the arts are just one manifestation of this. Evidence to support causal attribution is either lacking or weak in many areas of public policy, indeed this is the central problem in policy research, one highlighted by reviews and evaluations of the evidence base, often commissioned by government departments, in a wide range of areas (Davies et al., 2000). In the UK these include criminal justice (Harper & Chitty, 2005), community health (Bauld, Mackinnon, & Judge, 2001), public health (Millward et al., 2003), mental health (Wykes, 2004) and area-based regeneration (Burton et al., 2006).

Within the wider world of evaluation and social science research, this lack of evidence of causation has been linked to the limitations of the dominant “natural science” experimental model of research and evaluation. This view is based both on a pragmatic recognition of the difficulties of applying this model in the complex open systems of the “real world”, which in many contexts provides little opportunity for quasiexperimentation let alone a “gold standard” approach. It is also based upon an awareness of the ontological weakness of the successionist understanding of causal change that underpins the medical model and its limitations for handling the contextual complexity of many social interventions. As interventions (in our case, encounters with the arts) take place within the open systems of the social world, and as a result are almost never repeated in *exactly* the same conditions, this explains why experimental studies typically produce differing or conflicting findings, and consequently inconclusive evidence.³

Within the context of evidence based policy making and the “what matters is what works” agenda of the UK government since 1997, the perceived weaknesses of traditional evaluations based on this model – particularly the failure to produce clear evidence of impact – has led to a growing interest in alternative models of evaluation, and particularly theory-based evaluation (TBE) approaches (Pawson, 2006; Sanderson, 2002).

In contrast to the development of this thinking within wider public policy evaluation, in much of the discussion of the social impact of the arts there remains an implicit assumption both that causation *can* be definitively established and, that a medical, experimental model of research can deliver this (Cowling, 2004). The dominance of this model is reflected in its implicit acceptance by most (but certainly not all) commentators as the most appropriate one for investigating the social impacts of the arts, and therefore as the standard by which research evidence is assessed. This is apparent in many of the methodological criticisms, which tend to relate to concepts associated with this paradigm such as replication and comparability, generalization and representativeness. An example of this is Merli’s (2002) critique of the large scale empirical study of the social impact of the arts conducted by Comedia and reported in *Use or Ornament?* (Matarasso, 1997). Shortfalls identified by Merli include a research design that lacks control groups, failure to control for “extraneous variables” as required for causal attribution, and a sample of projects which is not representative of all types of participative arts activity in the country, and, therefore, findings that are not generalizable (p. 10).

The implicit acceptance of successionist logic is particularly evident in relation to case study based approaches to social impact, which are typically criticized for being both unrepresentative and ungeneralizable. The qualitative methods used are typically regarded as lacking robustness, in keeping with the hierarchy of evidence devised for use within medical research (an example being AEGIS, 2004, p. 53; Petticrew & Roberts, 2005). In contrast, other writers including Matarasso (1997), Jermyn (2004), and Belfiore and Bennett (2007, p. 233) have questioned the applicability and practicality of the natural science experimental approach for dealing with the complexity of arts impact. This includes a plea for the field of arts and mental health not to be forced into the “straitjacket of a medical model of research” (Macnaughton, White, & Stacy, 2005, p. 338).

Notably much of the literature about the social impact of the arts consists of reviews, critique or commentary on conceptual or methodological issues rather than empirical work.⁴ The main body of empirical studies of the effects of the arts on individuals had been conducted in fields such as medicine, psychology, gerontology, mental health and public health that share the dominant successionist view of causation. Following the critique of the medical model developed within wider public policy evaluation, it is therefore worth examining the argument that, as in other policy areas, many of the key deficiencies identified in the evidence base for the social impact of the arts are a consequence of the ontological limitations of this model and its successionist view of causation.⁵

Part 2: The arts and the successionist model of causation

Typically empirical studies of the arts attempt to measure the effects on individuals of some type of encounter with the arts, testing individuals before and after exposure to an arts experience. The types of effects to be measured and a way of measuring them are preselected. In psychology and health-related fields these are usually validated measures – either a single question or series of questions (a “scale”) – developed and tested for the purpose of measuring theoretical constructs such as subjective well-being, anxiety, self-confidence or quality of life. Statistical tests of association are run on the data and results reported.

In other words, such studies aim to establish with certainty whether an independent variable (x), i.e., arts participation, produces change in a dependent variable (y), i.e., a defined population. Through the use of a counterfactual in the form of a control group (i.e., the “gold standard” random control trial) and of rigorous randomized sampling, the purpose is to be able not only to establish with certainty whether a variable possesses causal powers, and to measure these, but to generalize the findings to a wider population. Some of the key criticisms of social impact studies relate to features of this simple successionist model of change; these are to do with generalizability, explanations of causal change and contextual complexity.

Generalizability

Many criticisms of impact studies are founded on successionist claims about the possibility of generalization. As we have seen above, with reference to the work of Belfiore and Bennett (2007), the literature on social impact of the arts has frequently questioned the presumption that findings about the effects of the arts, positive or negative, can be extrapolated to a wider population, and generalized claims made regarding how “all” arts will affect “all” people in “all” circumstances (Belfiore, 2006; Jermyn, 2001; Reeves, 2002, p. 102).

Yet the objective of studies in the natural science paradigm *is* to secure generalizable evidence of cause and effect in social phenomenon. Achieving this is largely a technical issue: the rigorous selection of a sample representative of the larger population.

The potential for generalization rests on another key assumption in successionist logic, that individuals play an essentially passive rather than active role in change. Part of the challenge that complexity and variance pose for the successionist view of causation is to do with differences in the motivations, interests and dispositions of individuals. One of the main findings of a review of studies of the contribution of arts and cultural participation to quality of life is the degree to which this varies according to the interests, background and circumstances of the individual concerned (Galloway, 2006). Coalter discusses a number of reasons why a direct relationship between arts participation and positive outcomes for individuals cannot be assured. These include differences in: the relative significance of the arts experience for individuals; the frequency and intensity of participation (i.e., the degree to which individuals engage, and how); and self-selection, or motivation – whether or not individuals have chosen to participate and the context in which the decision was made (2001, p. 10). All are issues of agency largely elided in practice by successionist studies.

Explanatory failings

A second type of criticism made of impact studies is that they overclaim the extent to which the arts may be responsible for observed effects, or that they do not explain the effects of the arts. This is a key epistemological limitation of studies based on the successionist model. They tell you whether an association exists (or not) between two variables, but can’t explain why, having to resort to post hoc explanations.⁶

An example from an arts impact context is Michalos' (2005) investigation of the impact of arts participation on the quality of life of residents in the Canadian city of Prince George. Respondents were asked about the frequency with which they took part in 66 arts-related activities,⁷ to rate the satisfaction they gained from each one, and to answer questions about their quality of life, using four different validated measures of self-perceived quality of life.⁸ Findings about the relationship between arts participation and quality of life varied according to the dependent variable (the particular measure of quality of life) used, and the significance of specific arts activities also differed accordingly. So, satisfaction with gourmet cooking and embroidery had the strongest positive association with life satisfaction. Satisfaction with attending cinema had the strongest correlation with global quality of life. Satisfaction with gourmet cooking and buying works of art was the strongest influence on happiness, whereas hours per week singing in a group was negatively associated with happiness. This example illustrates not just how problematic the question of measurement is within the social sciences, but also the problems of interpretation and explanation presented by a successionist model of change.

Complexity

A third key criticism of studies of social impact of the arts is that they often both overclaim and oversimplify the nature of the causal link between the arts and observed effects because they take insufficient account of the complexity of the contexts in which arts activity or encounters occur and, therefore, the wide range of other possible factors that may shape how the arts affect individuals (AEGIS, 2004; Reeves, 2002, p. 103). One particularly overlooked aspect of this complexity, say critics, is the issue of the aesthetic quality of the art being evaluated. It is argued that there is often an implicit assumption in evaluation or research work that all artforms, or all forms of arts activity or practice within an artform are the same (Belfiore, 2006; Hamilton & Scullion, 2006). Other writers including Matarasso (1996) and Coalter (2001) have argued for studies to take account of this complexity, including a focus on the attributes of arts projects which help explain their effects on participants, and the characteristics of those who take part.⁹

One of main reasons for this is that, in order to specify a causal relationship, successionist studies seek to control or suppress precisely those types of attributes which may play a role in determining the potential effects of the arts. These are purposively ruled out of the analysis either explicitly, by suppressing them, or implicitly, by making the assumption that these are constant.

The importance of this is underlined by recent work on the determinants of arts impact. Researchers in the social psychology of music have proposed a more complex model than the traditional "transmission" model to explain people's responses to music.¹⁰ According to this, the effects of music on individuals involves the mutual and complex interaction of three major determinants of impact, identified as the characteristics of the music; the characteristics of the person (with reference to the body of research evidence suggesting that people consciously *use* music to achieve certain psychological states); and the characteristics of the social situation in which the musical interaction occurs; they refer to a growing literature which shows how specific social and institutional contexts exert a powerful influence on people's responses to music (Hargreaves, MacDonald, & Miell, 2005). With reference to the much wider literature relating to the arts as a whole, Belfiore and Bennett conclude that social, cultural and psychological factors all combine to shape aesthetic experiences and they identify a very similar set of determinants of impact (2007).¹¹ They pose two key questions: (1) whether it is even possible to develop a methodology capable of investigating the intricacy and complexity of the aesthetic experience, and (2) given the specificity of this experience, "whether *any* meaningful generalizations can be made about the social impact of the arts" (2007, p. 262).

Part 3: Theory-based evaluation and the generative view of causation

Finding ways of dealing more effectively with these types of issues to do with specificity, context, complexity and agency, has been an important driver of interest in TBE, the growth of which is reflected in a burgeoning cross-disciplinary literature both advocating and critiquing the approach (Rogers, Petrosino, Huebner, & Hacsí, 2000, p. 6). Interest has been particularly strong in relation to large social interventions which present considerable challenges for evaluators (Pawson, 2006; Sanderson, 2002). Since 1997 TBE has been used in evaluations of major UK social policy initiatives including Health Action Zones, Sure Start and the Choose Life suicide prevention strategy.

Theory-based evaluation (TBE) has emerged in two main forms: in the UK, realistic evaluation (Pawson & Tilley, 1997) and in the USA, the Theory of Change approach (ToC) developed by the Aspen Institute (Fulbright-Anderson, Kubisch, & Connell, 1998). ToC was developed precisely to deal with the acknowledged difficulties of implementing RCT designs in large-scale complex community-based initiatives (CCIs). Meanwhile, realistic evaluation has been assessed as best serving individual-level change (Blamey & Mackenzie, 2007, p. 451; Connell & Kubisch, 1998).

This family of TBE approaches shares its ontological roots with realism within the social sciences (Bhaskar, 1975; Harre, 1972; Sayer, 1984). The key “offer” of TBE is its potential for strengthening causal attribution through new types of explanatory strategy based on a generative view of causation (Bickman, 1987; Chen, 1990; Rogers et al., 2000; Weiss, 1972, 1997). Implicit in TBE then is a critique of the successionist understanding of social change that underpins both the medical model of research and traditional forms of evaluation.

The generative model of change

A generative understanding of causation views change as attributable to the internal characteristics of objects and explained through the interaction between context, and mechanism or process (see Figure 1). Research based on a generative understanding of causal change sets out to understand: what is it about an intervention which might lead it to produce change, and crucially in what conditions, or sets of circumstances, is that change triggered. It is an approach best summed up in the phrase: what works, for whom, and in what circumstances? (Pawson & Tilley, 1997; Rogers et al., 2000).

Pawson and Tilley explain generative causation using the example of gunpowder (1997, p. 57). A chemical reaction occurs when a flame makes contact with gunpowder but *only if* it is dry, and if there is oxygen present (i.e., if the conditions are right). In other words, whether or not an explosion takes place is contingent not only on the internal attributes of the gunpowder (its chemical composition) but on the context in which a flame is applied (as well as on the strength of the flame and the duration of contact with it).

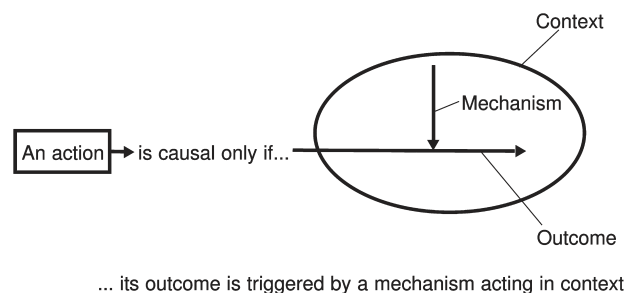


Figure 1. The generative model of causal change.
Source: Pawson and Tilley (1997, p. 58).

For research based on a generative model of change factors specific to particular contexts, whether a geographic area, a point in history or a type of institution, on which change may be contingent, are of core interest, rather than factors to be controlled for. As a result inquiry of this type is multilevel in focus (Sayer, 1984). And crucially a generative understanding views human beings as *agents* in social change, and change as occurring in open rather than closed systems.

This generative model has an intuitive appeal for arts impact research. First and most obviously, its analysis of causal change embraces all three types of determinant of impact identified in the literature, that is those factors inherent to the individual, the artwork and to the context, or environment (Belfiore & Bennett, 2007, p. 225; Landry et al., 1993; Matarasso, 1996; Miell, MacDonald, & Hargreaves, 2005). This includes consideration of the quality of aesthetic experience and artistic practice, for “it is the character of the contact that should be considered, not just how many contacts and of what duration” (Pawson & Tilley, 1997, p. 50). And as we shall see, it addresses the widely acknowledged problem of the lack of theory in relation to arts impact (AEGIS, 2004; Merli, 2002; Oakley, 2004).

Theory-based evaluation

Common features and differences between the two types

The shared features of theory-based forms of evaluation reflect their common ontological roots while the differences between realistic evaluation and ToC, which are beyond the scope of this discussion, have been summarized as follows:

they take different approaches to articulating theory, in practice generate different types of theory, require differing degrees of stakeholder involvement, are more or less suited to different degrees of complexity, and ultimately provide different types of knowledge for implementers, policy-makers and for use in future evaluations. (Blamey & Mackenzie, 2007, p. 452)

In practical terms realistic evaluation and ToC both involve, as a first step, the articulation of a theory or theories about how an intervention is intended to work. While there are differences in how the two approaches define and deal with theory, in both it is theory which informs the evaluation design and guides the collection and analysis of data. In realistic evaluation this theory is conceived of in terms of “propositions about how mechanisms are fired in contexts to produce outcomes”, represented in the form of “Context-Mechanism-Outcome” (C-M-O) configurations (Pawson & Tilley, 1997, p. 85). Meanwhile, ToC concentrates on identifying possible causal “paths”, based on a generative understanding of change, with the overall ToC comprising both a “program theory” and an “implementation theory” (Blamey & Mackenzie, 2007; Rogers et al., 2000; Weiss, 2000).¹²

Both approaches conceive of participants in interventions not as passive “subjects”, but as agents, whose “motivations, attitudes, previous knowledge, and skills are an important part of the context within which causal mechanisms work or fail to work” (Rogers et al., 2000, p. 51). Instead of producing “outcomes”, Pawson and Tilley argue that interventions should be seen as providing opportunities “which may (or may not) be triggered into action via the subject’s capacity to make choices” over whether or not to take part and how they choose to engage (1997, p. 38).

Theory-based evaluation and the arts

Evidence of theory-based approaches being adopted in the arts is relatively recent. Their influence is suggested in some US research developing cultural indicators of community quality of life.

A notable example is The Urban Institute's work on indicators which began with an intensive period of theory building over more than 2 years in which the definitions, meaning and value of community-based arts and cultural activity were explored with reference to existing research and data, and through extensive interviews and focus groups with "stakeholders" of all kinds in communities across the USA (Jackson, 1998; Jackson, Herranz, & Kabwasa-Green, 2003; Jackson, Kabwasa-Green, & Herranz, 2006). Equipped with this theory researchers have worked with a US-wide network of community arts organizations to develop tools to capture data relating to the theory-relevant dimensions of arts participation and impact. Another example is the Creative Community Framework of cultural indicators developed by Culture Initiatives Silicon Valley, which began with the articulation of a causal theory of the impact of the cultural sector on a community (2003, p. 5). In the UK there is evidence of logic modelling, a technique closely associated with (though not the same as) ToC being used within the cultural sector, for example, in the study of social impact in relation to museums (AEA Consulting, 2005).

Although TBE appears to offer an ontologically more robust research orientation for investigating arts impact, differences in types of TBE and the range of interpretations, particularly within ToC, makes it difficult to generalize about what TBE contributes practically to studies of the arts. The next section looks at this with reference to four recent studies of the social impact of the arts that adopt TBE. The discussion focuses on three main claims made for TBE: that it helps in developing theories of causal change; that it provides knowledge about how and why change occurs (what works, for whom, in what circumstances); and that it helps develop "evidence", particularly of causal attribution. In discussing the latter point we will consider how TBE addresses a particular problem for arts impact studies, which is the use of qualitative self-report data.

The aim is not to provide a detailed description of the aims, methods and findings of the four studies but to discuss the insights they provide about the utility of theory-based approaches for strengthening the evidence about arts impact. In the process the aim is to explicate more of what the approach involves, including its associated difficulties.¹³

Part 4: Theory-based approaches to the impact of the arts

The four studies

The four are UK studies of the arts in the context of social inclusion (Goodlad, Hamilton, & Taylor, 2002a), criminal justice (Miles & Clarke, 2006), mental health and social inclusion (Anglia Ruskin/UCLan, 2007a) and health (Kilroy, Garner, Parkinson, Kagan, & Senior, 2007a). Each adapts a theory-based approach for its own purpose, three using realistic evaluation and one ToC. They were commissioned by a combination of UK government departments, including the Department for Health, DCMS, Department for Education and Skills, HM Treasury and public bodies, including Arts Council England and the Scottish Arts Council. Table 1 provides an overview of the four studies, based on a matrix developed by Pawson and Tilley (1997, p. 88), which allows key features of the study designs to be compared.

Goodlad et al. (2002a) is a realistic evaluation of a national funding scheme encouraging the use of the arts as a means of tackling social exclusion in areas of multiple deprivation.¹⁴ The evaluation focuses on three dimensions of the scheme's effectiveness, including its impact on other agencies and partners; on public awareness; and on participants and the local community. Ten individual Social Inclusion Partnership (SIP) arts projects, most targeting young people, were selected for study. These represent a range of geographic areas, scale and duration of project, arts approaches and artforms, including music, drama, dance, crafts, writing, photography, art and design and video making.

Table 1. Matrix of realist study designs.

Topic	Anticipated impact	Unit of analysis	Research strategy	Theory development	Opportunities for experimental control	Techniques	Result
Social inclusion (Goodlad et al, 2002a) Realistic evaluation	Immediate benefits to individuals, long term impact on communities	Individual and community	Survey of SIPs and studies of 10 selected projects	Preliminary. Based on literature and expert knowledge	Limited. Focus group interview with Control groups of non-participants in each SIP	Interviews, focus groups, survey	Theory begun to be tested. Immediate impacts for participants, described.
Criminal justice (Miles & Clarke, 2006) Realistic evaluation	Immediate benefits to individuals (but primarily a study of research feasibility)	Individual	5 individual project evaluations enabling comparison by gender, artform & setting.	Based on review of literature, practice and theory (Hughes, 2004)	None, due to difficulties with phasing arts projects and participant recruitment	Profiling, psychometrics, observation, interviews, diaries, tracking Pre- and posttest measures	Theory begun to be tested. Some “indications of impact” established.
Mental health and social inclusion (Anglia Ruskin/ UCLan, 2007a) Theory of Change	Immediate	Individual	Realist case studies (qualitative) attached to quantitative experimental design	Commonalities in project level theories identified	None	Profiling Interviews with participants	Theory begun to be tested and refined Identified 8 distinct processes through which participants gained benefits
Health and well-being (Kilroy et al, 2007b) Realistic evaluation	Immediate	Individual	Evaluation across 6 individual projects	Preliminary	None	Pre and post test measures diaries interviews appreciative inquiry cycles	Theory begun to be tested and refined

Source: Adapted from Pawson and Tilley, Table 4.1 (1997, p. 88).

Miles and Clarke's study (2006) takes a realistic evaluation approach to investigating the effectiveness of the arts as a medium in offender rehabilitation.¹⁵ Commissioned as part of a strategy to improve the evidence base in arts and criminal justice, its main purpose is to explore the practical issues involved in evaluating arts in criminal justice projects and to test a range of data gathering methods including profiling, psychometrics, observation, interviews, diaries and tracking. The study embraces state and private sector establishments, projects delivered both in-house and by the voluntary sector, and some of the UK's leading arts and criminal justice organizations. The study involves five arts projects, mainly in custodial settings, using a combination of drama, dance and creative writing.

Anglia Ruskin/UCLan (2007a) use a ToC approach to investigate the "processes, benefits and outcomes" of arts participation for adults with mental health needs (2005; 2006a; 2006b; 2007a, p. 10; 2007b).¹⁶ It forms part of a research strategy designed to pilot approaches and methods for improving the evidence base for the arts. Indicators and measures for evaluating arts and mental health projects are developed and piloted. The outcome evaluation involves 88 participants in 22 arts projects across England. The research design comprises a traditional experimental pre- and posttest study, complemented by qualitative case studies of six selected arts projects. A ToC approach is adopted for the case study work, as a way of structuring and guiding an investigation of how, why and for whom arts initiatives work (or not) "through a systematic and cumulative study of the links between their activities, outcomes and contexts" (Anglia Ruskin/UCLan, 2007b, p. 3).

Realistic evaluation is used by the Invest to Save: Arts in Health Programme to understand the value and impact of arts participation on the health and well-being of three groups: older people, people with mental health needs, and health service workers (Kilroy et al., 2007a). As a partnership project it aims to meet a diverse set of "stakeholder" needs both for evidence of impact (health stakeholders) and understanding of the transformative effects of arts practice (arts stakeholders). The study also pilots participatory evaluation as a means of building evaluation capacity with practitioners from six arts in health projects. Practitioners were involved collaboratively in all stages of the research process. The study design involves pre- and posttest measurement of wellbeing, general health, depression and satisfaction using validated scales, and the use of diaries and interviews with participants.

Theory development

It is posited that theory based evaluation helps in articulating, testing and refining theories of causal change (Pawson & Tilley, 1997; Weiss, 2000). This is of huge relevance for the study of arts impact, an area generally characterized as lacking theoretical grounding and a particular criticism levelled at qualitative case study work (AEGIS, 2004, p. 58). The four studies reached the stage of either tentative theory development or of initial testing of theory with most commenting on the need for further studies and evaluations to build on this. Table 2 shows the process followed by the four studies as they seek to understand why and how interventions produce change.

Constraints of resources, timescales, scheduling and external pressures affect all evaluations whether traditional or theory based. Theory-based evaluation involves establishing criteria and then identifying, prioritizing and exploring the most promising CMOs (realistic evaluation) or program linkages (ToC) (Rogers et al., 2000). Constraints of time and scale, plus, in some cases, project cancellations and postponements and a multitude of practical constraints and barriers meant a systematic investigation of the optimum range of context-mechanism-outcome configurations was not possible (Goodlad, Hamilton, & Taylor, 2002b; Miles & Clarke, 2006, p. 4.1.1). For similar reasons the Invest to Save study prioritized one core theory comprising

Table 2. Four studies: mapping realist research process.

	Step 1 Understanding of the intervention, aims, context, settings, prevailing theories	Step 2 Mapping and prioritization of potential CMO configurations	Step 3 Outcome inquiry – building picture of the intervention in practice	Step 4 Theory refinement and development
Goodlad et al. (2002a) Realistic evaluation	Hypothesized how the scheme might contribute to social inclusion using available research evidence and experts' knowledge.	Elaborated a multidimensional model comprising possible pathways through which social inclusion might be assisted in the longer term (2002, p. 5).	Tested this model by means of interviews with SIP and arts project staff, survey work and focus groups of participants and nonparticipants.	Theory begun to be tested through identification of immediate impacts (longer term impacts outside the evaluation remit) and key contextual factors considered important for positive outcomes to occur.
Miles and Clarke (2006) Realistic evaluation	Hypothesized for each arts project how the intervention would lead to outcomes.	In projects where aims and objectives were clearly stated, two frameworks, resiliency theory and performance theory, were tested using tools designed for this purpose.	For each intervention (project) a combination of methods used to test its theory. Included profiling, psychometrics, observation, interviews, diaries, tracking.	Initial theory testing. Identification of intermediate outcomes, the building blocks for longer term change, a deeper understanding of the mechanisms involved in arts and criminal justice projects and the contexts in which they are most effective.
Anglia Ruskin/ UCLan (2007a) Theory of Change	Engaged stakeholders in each project in articulating a Theory of Change: how the activities would lead to outcomes.	Theories modified and refined using data gathered from participants. Identified 8 key processes that were triggered by the opportunities provided by the arts projects.	Used logical framework analysis of interview transcripts to begin developing an understanding of which processes were most important for whom.	Preliminary development/ refinement of theory for testing in a larger scale study. Complements the quantitative experimental aspect of the study which measured impact using validated scales.
Kilroy et al. (2007a) Realistic evaluation	Appreciative Inquiry workshops with stakeholders, practitioners and participants, developed "statement of intent" for each project.	Identified core process common to all 6 projects for prioritization in investigation.	Semistructured interviews with participants, artists and stakeholders used to elicit insights and test theory.	Refining of theory into a proposed "model of transformational change" identifying contextual factors necessary for longer term change.

the processes common to the arts projects involved in their study (Kilroy et al., 2007a). But as we shall see later, the emphasis on cumulative learning in theory-based evaluation still enables studies to contribute towards the bigger picture of theory development, either at a programme level (ToC) or within specific subject/topic areas (realistic evaluation).

The potential of the approach for theory building is demonstrated well by the Anglia Ruskin/UCLan “Mental Health, Social Inclusion and Arts” case studies. The first stage of these comprised a series of facilitated workshops with project representatives to explore understandings and views of how the projects worked and why, using the notion of participants’ “journeys”. What emerged from this process was not a generic ToC for arts projects, but a range of complex project-level “theories of change”, related to three “key outcomes” (Anglia Ruskin/UCLan, 2007b, p. 11). These theories were tested using data gathered during in-depth interviews with arts participants from each of the projects and both refined and developed, as in some cases processes triggered by participants’ experiences “were not so well explained by the projects’ own theories” (p. 11).

Ultimately the study identified eight clear but interrelated processes triggered by arts and mental health projects, from which individuals associated immediate benefits from taking part. The importance of these processes was observed to vary across projects “depending on the project’s context, the participants they aimed to reach and their particular approach”. The processes important to participants in *all* projects were: getting motivated; focusing; and connecting with others. Processes identified as very important in *some* projects were: self expression; connecting with abilities and having time out; and two processes, rebuilding identities and expanding horizons, were found to be important for *some* participants in *all* projects (Anglia Ruskin/UCLan, 2007b, p. 62). In recent work the authors have further developed a ToC for arts in mental health projects by linking what was learned from the study back to theoretical work on the social notion of “recovery” (Anglia Ruskin/UCLan, 2007b; Spandler, Secker, Kent, Hacking, & Shenton, 2007).

The other studies test theory in different ways. For example, Goodlad et al. identify five cultural and social contextual factors they consider key to explaining how an arts and social inclusion funding scheme effects change. The factors, or “key interactions”, found most important for providing a “context for change” include effective local staff, working with groups, achieving quality in the arts and developing effective partnerships (2002b, p. 9). They provide some evidence of their explanatory power, drawing useful lessons for future arts programmes (p. 14). The authors conclude that they have made “some progress” in suggesting the CMO configurations “most likely to achieve the desired effect of social inclusion through the arts” but add “there is a need for further evaluations to develop our understandings of how these operate” (p. 9).

While explicating the complexity of contextual factors at play in specific circumstances, Miles and Clarke’s study draws some general conclusions across the five arts projects. Evidence synthesized from the range of methods piloted demonstrates the potential of arts interventions to contribute towards greater self-esteem, self-control, confidence and co-operation amongst participants. In their view “what works” about arts participation in a criminal justice setting relates to “the arts process as a framework for and method of engagement rather than the specific context of one or other particular artform” (2006, Pt. 7). What participants gain is “respect and responsibility in a disciplined and creatively challenging, yet safe, supportive and crucially non-judgmental space ...” (Pt. 7). As with other studies here, the mechanisms produced varying outcomes dependent on context; the attributes of participating individuals, the social circumstances and the specific arts project. However in these custodial settings the authors suggest the greatest impact is felt by younger, more vulnerable individuals (Pt. 7). They emphasize that demonstrating the sustainability and transferability of the immediate impacts observed requires “longer-term or sequential projects supported by longitudinal research and an effective tracking methodology” (Pt. 7).

The outcome of the Invest to Save study is a proposed “model of transformational change” demonstrating how outcomes beneficial for health and wellbeing may be possible through the mechanism of arts engagement. Like the Anglia Ruskin/UCLan study the authors identify a number of processes deriving from arts participation, described as forgetting/distraction; challenge, potential for experimentation, support, shared experience, and a nonhierarchical and noncompetitive space in which to “practice being well”. The model also builds in the contextual attributes in which these are most likely to be “fired” (Kilroy et al., 2007b).

Articulation of theory

The starting point for TBE is the articulation of theory(ies) and this, particularly within ToC, requires investment in what can be a time consuming process, requiring evaluators with a specific skillset and necessitating a shift in the traditional detached role of researcher/evaluator (Blamey & Mackenzie, 2007). Although theory generation need not depend on it exclusively, an important part of articulating the theory(ies) of interventions in an arts context involves the communication by arts practitioners of how their practice contributes towards intended outcomes. In practice this is not straightforward. There is often wariness amongst practitioners of all types, because of political context and concerns about performance measurement, to setting detailed outcome targets and indicators (Mackenzie & Benzeval, 2005). In the arts there are particular sensitivities that reflect both widely varying models of arts practice and fundamental differences in perspective about whether the potential social impacts of the arts are the aim of, or incidental to, arts practice (Matarasso, 1996, p. 18).¹⁷ In addition to this are concerns that the prescribing of outcomes, and the potential invasiveness of research and evaluation methods will interfere with the flexibility and responsiveness necessary to arts practice (Miles & Clarke, 2006, p. 4.1.2; Shaw, 1999).

These issues are compounded where ToC is used, because the initial aim of the approach is for the theory to be “owned” by all stakeholders and consequently for an initial consensus to be achieved on aims, objectives and outcomes. However, as we know, competing stakeholder perspectives and needs are a common problem in evaluation work (Barnes et al., 2005; Martin & Sanderson, 1999) and this is a well-documented “thorny issue” in arts evaluation, where arts practitioners and their project partners/funders often have not only very different perspectives but also lack a common language with which to communicate (Coalter, 2001; Cowling, 2004, p. 128; Jermyn, 2001; Matarasso, 1996; McDonnell & Shellard, 2006; Reeves, 2002, p. 36).

This is borne out in the studies. Miles and Clarke report that in some cases they “had to second guess anticipated impacts because aims and objectives were only partially communicated or revealed at the last minute” (2006, p. 4.2.1) and that weaknesses in the type of programme theory they were able to generate, constrained the way in which this was able to guide method (p. 5.1). As a result, two of the studies (Goodlad et al., 2002b; Miles & Clarke, 2006) relied more upon existing literature and expert knowledge for theory generation rather than what Pawson and Tilley (1997) call the “folk theory” of practitioners. For example, Goodlad et al. drew on the wider academic and evaluation literature about urban policy, area-based regeneration and social inclusion, whereas Miles and Clark drew on a review of theoretical frameworks employed in arts and criminal justice field, selecting two specific theoretical models for testing: resiliency theory and performance theory (Hughes, 2004).¹⁸

The synthesis of realistic and participatory evaluation adopted by the Invest to Save study, combined with its use of Appreciative Inquiry, appeared to allow it to successfully negotiate a way through the initial wariness, tapping into the expertise and experience of arts practitioners, and producing the insights needed to surface theory.¹⁹ A key part of this was working to develop a “common language” not only between practitioners and stakeholders but also

between the researchers and the volunteers and practitioners in the arts projects (Kilroy et al., 2007a, 2007b).

Critical deconstruction

Theory has the “power to illuminate and to really shed light, just new light and to look at the world in a very different way [than] . . . we would without that kind of approach.” (ToC research practitioner)

Burton et al. talk about “surfacing” the theories underlying interventions that usually remain unarticulated and unscrutinized (2006, p. 309). This resonates for the arts, where many criticisms have been levelled at the assumptions on which policy are based (Belfiore, 2002; Oakley, 2004). Goodlad et al. (2002b) provide an example of how the reflexivity of a theory-based approach enables a critical deconstruction of policy interventions. Through elaborating and initially testing a theory, setting out the mechanisms through which arts participation might plausibly contribute to social inclusion, the authors conclude that “. . . at best, social inclusion *might* be supported by participation in the arts, but it is *unlikely* to be the *only* factor in securing inclusion in more than a very few distinctive cases” (Goodlad et al., 2002b, p. 14, emphasis added). This conclusion carries an implicit criticism of the claims made for the arts. A strength of the approach is therefore its reflexive approach both to theory and practice, combined with a concern for context, that mitigates against the overclaiming of impacts (Weiss, 1998).

Evidence: causal attribution and wider issues

Based on interviews with arts and cultural policy staff in both the UK and Australia, the Australian Expert Group on Industry Studies reports “confusion about what constitutes evidence and the level of rigor needed to justify focusing funds on programmes with specific social impact aims” (AEGIS, 2004, p. 15). As argued earlier, notions of what constitutes “evidence” and of a hierarchy of evidence are strongly influenced by the medical model of research (Petticrew & Roberts, 2005). Theory-based evaluation rejects this and instead deals with the problem of evidence in a number of ways.

First, on the key question of causation, it is argued that TBE, particularly ToC, provides policy makers with a convincing “narrative framework” or “explanatory synthesis”, in which a range of types of evidence can be mobilized to support causal attribution. A distinctive feature of TBE (and realist social science) is its methodological pluralism. This includes the use of experimentation, of a realist kind, based on a generative model of change (Pawson & Tilley, 1997). Miles and Clarke, in their study, highlight this as a key strength, arguing that “the adoption of a multi-method realist evaluation framework offers the arts in criminal justice a more sympathetic and workable research model, which arguably, is more powerful when it comes to distilling and explaining intervention impacts” (2006, p. 5).

Second, ToC is often used as a tool for planning the implementation of interventions with the programme theory used to determine indicators of change at end- and mid-points. Because it is concerned with mechanisms and contexts and examining the links between these, it allows immediate observed effects to be interpreted and linked to potential longer term impacts which lie beyond the evaluation period. Within the framework of a strong theory, the presence or absence of a phenomenon at a mid-point is therefore viewed as a reasonable predictor of longer term outcomes. If applied to major culture-led regeneration schemes, for example, ToC might help with the problems of lack of evidence of impact identified by Evans (2005, p. 960), which, in his view, are “in part due to problems of measurement and evaluation criteria not being established a priori or consensually, but also to the fact that, like hallmark event projects,

major culture led regeneration schemes are not wholly grounded or rational-decision-based” (p. 960).

In ToC the level of evidence for causal attribution is pragmatically determined. Blamey and Mackenzie explain that ToC “leaves the door open for the stakeholder group to determine the level of agreement that is required between its prospectively agreed theory and its outcomes in practice”, so stakeholders may be satisfied that it “works” (even if outsiders are sceptical) (2007, p. 449). The Invest to Save study appears to exemplify this, its synthesis of participative and realistic evaluation producing something akin to ToC in practice. In this study, “stakeholders” from both the arts and health fields were involved in developing projects’ theories of change and the evidence suggests that their shared “ownership” of theory did indeed help achieve both a common sense of purpose and focus on the research “journey”, and a convincing standard of evidence for stakeholders (Kilroy et al., 2007a).²⁰

It is really, really useful to have a strong underpinning theory driving it at the core that everyone can identify with . . . if the theory is driving it, everyone can respond to that. (TBE research practitioner)

In contrast realistic evaluation is much more focused on causation. One key objection is the potential for theory based inquiry to be “self-confirming” (Sanderson, 2000). To avoid this, and therefore strengthen the standard of evidence, various strategies are proposed to build counterfactuals into the investigation, chiefly through a methodological focus on identifying “outliers”, exceptions, contradictions, and on explaining unusual cases. Realist evaluators therefore need to maintain a focus on the fact that actual impacts may be counterintuitive and also adverse (Sanderson, 2000; Thomson et al., 2006) and unintended outcomes are as important to identify as intended (Chen, 1990). For practitioners of ToC it requires “that they develop strong theories, use multiple methods of enquiry to search for and confirm patterns in data, creatively blend designs and refrain from rushing to judgement based on findings from individual studies” (Granger, 1998, p. 1).

The problem of “anecdotal” evidence

The realist method has a number of advantages, not least of which . . . is the way it makes a virtue of the stories and insights that those most concerned with the projects can tell. (Goodlad et al., 2002b, p. 15)

In traditional social science and evaluation approaches, confidence in findings depends in large measure on demonstrating the validity or “integrity” of data. In an arts context self-report data gathered through methods such as interviews, focus groups or self-completion questionnaires is generally viewed as a weak (“anecdotal”) form of evidence. Within a realist framework, data is viewed in a way that effectively eliminates this issue, for in TBE approaches data is first and foremost a way of confirming, falsifying or refining theory, and the self-report of participants (or their family members, friends, carers, teachers etc.) is just one of many possible ways explored for testing theory. An example are the Anglia Ruskin/UCLan case studies where participants’ testimony is rigorously interrogated in order to test practitioners’ views and academic theory and ultimately to refine projects’ ToC (2007b).

For realist investigators, the testimony of those actually involved in an artistic encounter (or other “intervention”) allows unrivalled insights into individual motivations, choices, circumstances and perspectives (Miles & Clarke, 2006, p. 5.4; Pawson & Tilley, 1997, p. 162). Across the four studies this was singled out as a key strength of a realist approach, another being the way in which a realist or TBE approach provides a coherent structure for the collection, interpretation

and analysis of self-report data. The common perception of self-report evidence as weak in traditional data-driven evaluations is largely due to the lack of an explicit methodological framework for data analysis and a lack of transparency about the selection of quotes (Corden & Sainsbury, 2006). As we have seen, this is a particular criticism made of studies of arts impact using primarily qualitative methods. This is explicitly addressed by theory-based approaches, as Miles and Clarke demonstrate in their insight into the method of realist interviews used in their study of arts in criminal justice:

Interview schedules were initially built up from . . . applicable theoretical frames and usually based on . . . a particular project's aims and the types and contexts of potential effects. This recursive approach (a) allows reflection back on the utility of a particular theory and therefore contributes to theory building and (b) encourages personal narratives to be articulated and understood. (2006, p. 5.4.3)

Discussion

In summary the four studies illustrate how a theory-based approach can help address in a practical way some of the central issues in the discussion of the social impact of the arts. However they also illustrate some limitations, widely discussed in the evaluation literature (Blamey & Mackenzie, 2007; Sanderson, 2002).

Compared with the traditional experimental or quasiexperimental successionist model, a TBE approach provides useful insights into how and why engagement with the arts might effect changes in individuals, both positive and negative. The question is whether TBE need necessarily be viewed in opposition to traditional experimental approaches, or whether the two can complement each other (Cook, 2000). This is the subject of debate in the TBE literature and the four studies reflect differing views. In terms of practicalities the studies highlight particular difficulties and barriers in an arts context of articulating theory(ies) of change; identifying the most suitable or strongest theory to test; and of planning and implementing a realist research design capable of testing theory rigorously.

The clear advantage of TBE is that it draws upon all of the acknowledged determinants of change within a generative understanding. However, there are problems. One is that arts contexts themselves present difficulties for research feasibility, as much for theory-based approaches as traditional ones. The fragile economy of arts in mental health and arts in health work is well documented in the group of studies discussed (Anglia Ruskin/UCLan, 2007a, p. 9; Kilroy et al., 2007b). If selection of case studies and recruitment of participants is directed by the desire to investigate specific Context-Mechanism-Outcome configurations then the logistical possibilities for doing so are constrained.

Second, in practical terms a single study using a TBE approach is constrained in terms of the degree of complexity it can accommodate; each of the four studies had to narrow its focus, illustrating that handling complexity is not just an issue for traditional experimental approaches. Consequently as a group the four studies each contribute useful insights into how and why different types of change might occur for different individuals, but in terms of theory development the advances made are modest. This is worth discussing in more detail because it relates to one of the key methodological issues raised in the social impact of the arts literature, on which TBE has something important to contribute.

This is how to maximize the value of case study work that is usually small scale, takes place in very specific contexts, and uses predominantly qualitative methods, an issue that is of wider concern within social research and evidence based policy making (Britten et al., 2002; Spencer, Ritchie, Lewis, & Dillon, 2003; Thomas et al., 2004). Guetzkow (2002) amongst others has argued the need for a method and way of grouping, for evaluation purposes, large

numbers of similar studies in order to be able to draw out generalized conclusions (Burns Owen Partnership, 2005; Evans & Shaw, 2004; White, 2004).

TBE approaches eschew the concept of generalizability in favour of “cumulation”, the idea that, drawing on a body of primary studies, we can build knowledge about how and why change occurs. ToC is concerned with this at the level of large-scale programmes of complex community interventions (CCIs). With a strong theory evaluators can anticipate and examine how differences in context may alter effects. If we understand how and why interventions produce effects for some individuals (but not others) in some contexts then this knowledge might possibly be applied more widely, although not through “mechanical replication” (Weiss, 2000).

Pawson and Tilley state the realistic evaluation position on cumulation thus:

Rather than anticipating the cumulation of program wisdom in the form of discovering representative programs which work universally, the realist evaluator seeks to generalise about programs through a process of CMO configuration abstraction, the creation of middle range theories which provide analytic frameworks to interpret similarities and differences between families of programs. (1997, p. 217)

In practice this involves planning a series of studies enabling a single mechanism to be tested in a range of contexts. Or, alternatively, by identifying the Context-Mechanism-Outcome configurations contained in *existing* evaluations or studies and using these, retrospectively, to test theory. In effect what is proposed is a realist form of systematic review, in which in contrast to traditional reviews, the unit of analysis is the theory to be tested rather than the individual study (Pawson & Tilley, 1997, p. 152). Pawson’s recent work presents his recommendations for a realist method of research synthesis (2006).²¹

Meanwhile the ideas of TBE have also influenced thinking within the positivist paradigm, as evidenced in developing methodological approaches to meta-analysis, in which the capture of data about context and mechanism (process) are regarded as essential (Petticrew & Roberts, 2005).²² Indeed Winner and Cooper (2000) are early advocates of this in relation to the meta-analysis of arts impacts. What these realist or realist-influenced approaches offer is a way of overcoming the limitations of the traditional narrative review methods deployed within arts impact and an entirely different and potentially illuminating way of drawing out the lessons of empirical studies and using them to develop a theory of impact.

Part 5: Conclusions

The assumption, usually implicit, that the medical model of research with its “hierarchy” of evidence, and standards of proof, is appropriate or useful in all research areas is pervasive. The appeal and influence of this model is not difficult to understand: it appears, to some, to offer certainty, “scientifically” objective and definitive answers, and methodological “robustness”. Definitive proof of causality is elusive. Yet there is surprisingly little acknowledgement of this in critiques of the social impact of the arts. Nor is there much discussion of the types and standards of proof of causal attribution that are appropriate for the arts or cultural policy making, bearing in mind that these vary according to context (Davidson, 2000, p. 23).

TBE (and realist social research), based on a generative view of change, are ontologically best suited to studying the potential social impacts of the arts. A key strength of TBE approaches, at a theoretical level, is the way they speak to some of the central problems identified in the research evidence for the social impacts of the arts (while acknowledging that ToC and realistic evaluation often differ in how they do so). By virtue of the very different perspective they take, TBE approaches can highlight unarticulated assumptions and values, both about the arts and about

types of evidence.²³ The group of studies discussed here made small steps towards developing the theories of causal change claimed for TBE, but what they do illustrate clearly is the advantage of multiple and varied research methods for constructing convincing evidence. The particular types of knowledge produced by forms of TBE, about how and why change happens (or not), inherently challenges some of the more dogmatic expectations of the arts within the current target-driven style of public management.

Despite the ubiquitous calls, the political likelihood and ethical justification for investing substantial resources in large-scale longitudinal evaluations remains slim. The expectation of achieving a “once and for all” answer from one large study is also problematic. A key question remains how best to learn from the aggregation of smaller studies. In this quest a realist approach offers the potential for overcoming the limitations of both narrative and medical systematic forms of review. Much useful data already exists from studies in many areas of the arts. Approaches such as realist synthesis and realist influenced meta-analysis offer a way forward for developing both evidence and theory, feeding into future evaluation work, and should be recommended in this area.

Acknowledgements

The author would like to acknowledge the support of the Arts and Humanities Research Council and to thank the anonymous referees and colleagues who provided helpful suggestions.

Notes

1. For example Reeves (2002) summarizes the range of reasons for the deficiencies in social impact evaluations identified by Blake Stevenson Ltd (2000), Jermyn (2001), Matarasso (1996), Moriarty (1997) and Shaw (1999).
2. The reasons ascribed to this are remarkably familiar to the arts. They include the lack of co-ordination and incentives for academic research in this area (Millward et al., 2003, p. 31), a point referred to in relation to the arts (Belfiore, 2002; Oakley, 2004; Selwood, 2006).
3. In the arts, for example, this has been found to be the case for studies looking at the relationship with individual quality of life (Galloway, 2006) and academic attainment (Winner & Cooper, 2000). The latter involved a rigorous meta-analysis of studies which found a picture of conflicting findings, particularly between studies carried out in different countries which the authors felt might be accounted for by national cultural differences influencing the status of the arts within school education and the context, and therefore motivations, for pupils to choose to study the arts (p. 63).
4. For example, White and Rentschler’s literature review identified 12 key articles, four of which were empirical studies and eight of which were conceptual (2006, p. 4).
5. The few explicit criticisms of successionism made in the context of the arts include Matarasso, who discussed the relative strengths and weaknesses of causal link and causal mechanism analysis in his 1996 Working Paper for Comedia (1996, pp. 19–20) and, more recently, Miles and Clarke in the context of the arts in criminal justice (2006). Implicitly, the main reason for Belfiore and Bennett’s negative assessment of empirical research relates to the successionist model of causation deployed, with its “view of perception as a merely passive process (in which the percipient merely responds to the visual, musical or linguistic stimuli)” (2007, p. 238).
6. In evaluation these are typically concerned with three issues: did the intervention fail to produce the predicted outcomes, or is it the types of delivery of the intervention which failed, or did defects in the evaluation methods used fail to identify or measure the outcomes? (Rychetnik, Frommer, Hawe, & Shiell, 2002).
7. The definition used was developed and agreed in partnership with the local Community Arts Council.
8. These were overall life satisfaction, an aggregate of satisfaction with each of a number of specific quality of life domains, happiness and subjective well-being.
9. In fact these are some of a range of factors outlined by Matarasso as likely to influence the outcomes of an arts encounter. He also mentions the location and social situation of the art work and the resources available; the aims, agenda and relationships of stakeholders; the personalities, skills and experience of the arts practitioners involved; and the practicalities of timescale etc. (1996, p. 21).

10. In which a musical stimulus acts on a (passive) listener to produce and emotional, cognitive or behavioural response.
11. They propose three groupings of determinants of impact: factors intrinsic to individuals, factors intrinsic to the artwork or arts activity, and factors intrinsic to the social context in which the arts encounter takes place (Belfiore & Bennett, 2007).
12. The latter aims to make explicit the connections between interventions and outcomes, and so provide a plausible explanation of how and why the desired change is expected to come about. Weiss argues that, in practice, ToC has been little used to address causal mechanisms, tending to be concerned with theories of programme implementation, in other words, the links between the intervention activities and outcomes (Weiss, 1997). For the realistic evaluation perspective on types of theory in ToC see Pawson and Tilley (1997, p. 26).
13. Although the studies are categorized here as examples of TBE, it should be noted that the authors themselves refer sometimes to their methods as realist and, subsequently, the terms are sometimes used interchangeably here.
14. The Scottish Arts Council's National Lottery Arts in Social Inclusion scheme, which is targeted at Social Inclusion Partnerships (SIPs).
15. The study was commissioned by Arts Council England, the Department for Culture, Media and Sport and the Offender's Learning and Skills Unit at the Department for Education and Skills and forms part of a research plan initiated by the Research into Arts and Criminal Justice Think Tank (REACTT).
16. The study is significant first, because it represents the first large scale arts and mental health evaluation to be conducted in the UK, and second, because it applies validated measures of mental health outcomes which have tended not to be used in arts and mental health evaluations (Anglia Ruskin/UCLan, 2005, p. 4).
17. According to the latter view arts practice is free or separate from context. For an elucidation of the intellectual origins of this division, see Belfiore and Bennett (2006, p. 20).
18. These theories were selected because, in the author's view, they appeared to have the most scope for applicability to the selected projects within the constraints of the study (Miles & Clarke, 2006, p. 5.1).
19. For a description of Appreciate Inquiry, a philosophy of affirmative change management, see Kilroy et al. (2007a, p. 20) or Hammond and Royal (2001).
20. Information provided to the author in a personal interview.
21. Examples of realist approaches to this question within the social sciences include Ragin (1987).
22. It is worth noting that, in contrast to Pawson's position, other practitioners of TBE view it as compatible with traditional experimental approaches (see Cook, 2000).
23. In this group of studies evidence of "hard" outcomes was desired based on assumptions made, usually about vulnerable groups, for example that paid employment, less medication or more "integration" with mainstream society is necessarily beneficial.

References

- AEA Consulting. (2005). *Tyne and Wear Museums, Bristol's Museums, Galleries & Archives. Social impact programme assessment*. Museums Libraries Archives North East. Retrieved April 1, 2009, from <http://www.twmuseums.org.uk>
- Anglia Ruskin/UCLan Research Team. (2005). *Mental health, social inclusion and the arts: Developing the evidence base. Final report from Phase 1: The state of the art in England*. National Social Inclusion Programme. Retrieved April 1, 2009, from <http://www.socialinclusion.org.uk>
- Anglia Ruskin/UCLan Research Team. (2006a). *Mental health, social inclusion and arts: Developing the evidence base. Interim report from Phase 2: Retrospective analysis of project data*. National Social Inclusion Programme. Retrieved April 1, 2009, from <http://www.socialinclusion.org.uk>
- Anglia Ruskin/UCLan Research Team. (2006b). *Mental health, social inclusion and the arts. Second interim report from Phase Two: Baseline results*. National Social Inclusion Programme. Retrieved April 1, 2009, from <http://www.socialinclusion.org.uk>
- Anglia Ruskin/UCLan Research Team. (2007a). *Mental health, social inclusion and the arts: Developing the evidence base. Final report*. National Social Exclusion Programme. Retrieved April 1, 2009, from <http://www.socialinclusion.org.uk>
- Anglia Ruskin/UCLan Research Team. (2007b). *Case studies: Exploring the processes, benefits and outcomes of arts participation*. Unpublished report from the study "Mental Health, Social Inclusion and Arts: Developing the Evidence Base". Available from hspandler@uclan.ac.uk

- Australian Expert Group in Industry Studies of the University of Western Australia. (2004). *Social impacts of participation in the arts and cultural activities. Stage Two report: Evidence, issue and recommendations*. Cultural Ministers Council, Statistics Working Group (Australia). Retrieved April 1, 2009, from <http://www.culturaldata.gov.au>
- Barnes, M., Bauld, L., Benzeval, M., Judge, K., Mackenzie, M., & Sullivan, H. (2005). *Health Action Zones: Partnerships for health equity*. Abingdon, UK: Routledge.
- Bauld, L., Mackinnon, J., & Judge, K. (2001). *Community health initiatives: Recent policy developments and the emerging evidence base*. London: Neighbourhood Renewal Unit, Office of the Deputy Prime Minister.
- Belfiore, E. (2002). Art as a means towards alleviating social exclusion: Does it really work? – A critique of instrumental cultural policies and social impact studies in the UK. *International Journal of Cultural Policy*, 8(1), 91–106.
- Belfiore, E. (2006). The social impacts of the arts – myth or reality? In M. Mirza (Ed.), *Culture vultures: Is UK arts policy damaging the arts?* (pp. 20–37). London: Policy Exchange.
- Belfiore, E., & Bennett, O. (2006). *Rethinking the social impact of the arts: A critical-historical review* (Research Paper No. 9). Warwick, UK: Centre for Cultural Policy Studies, University of Warwick.
- Belfiore, E., & Bennett, O. (2007). Determinants of impact: Towards a better understanding of encounters with the arts. *Cultural Trends*, 16, 225–275.
- Bhaskar, R. (1975). *A realist theory of science*. Leeds, UK: Leeds Books.
- Bickman, L. (1987). The functions of program theory. *New Directions for Program Evaluation*, 33(Spring), 5–18.
- Blackman, T., Greene, A., Hunter, D.J., McKee, L., Elliott, E., Harrington, B., et al. (2006). Performance assessment and wicked problems: The case of health inequalities. *Public Policy and Administration*, 21, 66–80.
- Blake Stevenson Limited. (2000). *The role of the arts in regeneration*. Edinburgh, UK: Scottish Executive Central Research Unit.
- Blamey, A., & Mackenzie, M. (2007). Theories of change and realistic evaluation: Peas in a pod or apples and oranges? *Evaluation*, 13, 439–455.
- Britten, N., Campbell, R., Pope, C., Donovan, J., Morgan, M., & Pill, R. (2002). Using meta ethnography to synthesise qualitative research: A worked example. *Journal of Health Services Research and Policy*, 7, 209–215.
- Brown, S. (2005). How can research inform ideas of good practice in teaching? The contribution of some official initiatives in the UK. *Cambridge Journal of Education*, 35(3), 383–405.
- Burns Owen Partnership. (2005). *New directions in social policy: Developing the evidence base for museums, libraries and archives in England*. London: Museums Libraries Archives Council.
- Burton, P., Goodlad, R., & Croft, J. (2006). How would we know what works? Context and complexity in the evaluation of community involvement. *Evaluation*, 12, 294–312.
- Chen, H.T. (1990). *Theory-driven evaluations*. Thousand Oaks, CA: Sage.
- Coalter, F. (2001). *Realising the potential of cultural services: The case for the arts* (Research Briefing 12.43). London: Local Government Association.
- Connell, J.P., & Kubisch, A.C. (1998). Applying a theory of change approach to the evaluation of comprehensive community initiatives: Progress, prospects and problems. In K. Fulbright-Anderson, A.C. Kubisch, & J.P. Connell (Eds.), *New approaches to evaluating community initiatives: Vol. 2. Theory, measurement and analysis*. The Aspen Institute. Retrieved April 27, 2007, from <http://www.aspeninstitute.org>
- Cook, T.D. (2000). The false choice between theory-based evaluation and experimentation. *New Directions for Evaluation*, 87, 27–34.
- Corden, A., & Sainsbury, R. (2006). *Using verbatim quotations in reporting qualitative social research: Researchers' views*. York, UK: Social Policy Research Unit, University of York.
- Cowling, J. (2004). *For art's sake? Society and the arts in the 21st century*. London: Institute for Public Policy Research.
- Cultural Initiatives Silicon Valley. (2003). *Creative community index: Measuring progress toward a vibrant Silicon Valley*. Silicon Valley, CA: Cultural Initiatives Silicon Valley.
- Davidson, E.J. (2000). Ascertaining causality in theory-based evaluation. *New Directions in Evaluation*, 87, 17–26.
- Davies, H., Nutley, S., & Smith, P. (2000). *What works? Evidence-based policy and practice in public services*. Bristol, UK: Policy Press.

- Evans, G. (2005). Measure for measure: Evaluating the evidence of culture's contribution to regeneration. *Urban Studies*, 42, 959–983.
- Evans, G., & Shaw, P. (2004). *The contribution of culture to regeneration in the UK: A review of evidence*. London: Department of Culture Media and Sport.
- Fulbright-Anderson, K., Kubisch, A.C., & Connell, J.P. (1998). *New approaches to evaluating community initiatives: Vol. 2. Theory, measurement and analysis*. The Aspen Institute. Retrieved April 27, 2007, from <http://www.aspeninstitute.org>
- Galloway, S. (2006). *Well-being and quality of life: Measuring the benefits of culture and sport. A literature review*. Edinburgh, UK: Scottish Executive.
- Goodlad, R., Hamilton, C., & Taylor, P.D. (2002a). *Not just a treat: Arts and social inclusion. A report to the Scottish Arts Council*. Edinburgh, UK: Scottish Arts Council.
- Goodlad, R., Hamilton, C., & Taylor, P.D. (2002b, December). "Not just a treat": Issues in evaluating arts programmes to secure social inclusion. Paper presented at the UK Evaluation Society conference, "The Art of Evaluation: Artistry, Discipline and Delivery", London. UK Evaluation Society. Retrieved June 11, 2007, from <http://www.evaluation.org.uk/conference/Conf%20presentations%202002/Goodlad%20Hamilton%20&%20Taylor.pdf>
- Granger, R.C. (1998). Establishing causality in evaluations of comprehensive community initiatives. In K. Fulbright-Anderson, A.C. Kubisch, & J.P. Connell (Eds.), *New approaches to evaluating community initiatives: Vol. 2. theory, measurement and analysis*. The Aspen Institute. Retrieved April 27, 2007, from <http://www.aspeninstitute.org>
- Guetzkow, J. (2002). *How the arts impact communities: An introduction to the literature on arts impact studies* (Working Paper Series No. 20). Princeton, NJ: Princeton University Center for Arts and Cultural Policy Studies.
- Hamilton, C., & Scullion, A. (2006). *Quality of life and wellbeing: Measuring the benefits of culture and sport: A thinkpiece*. Scottish Executive. Retrieved April 1, 2009, from <http://www.scotland.gov.uk>
- Hammond, S.A., & Royal, C. (2001). *Lessons from the field: Applying appreciative enquiry*. Bend, OR: Think Book Publishing.
- Hargreaves, D.J., MacDonald, R.A.R., & Miell, D. (2005). How do people communicate using music? In D. Miell, R.A.R. MacDonald, & D.J. Hargreaves (Eds.), *Musical communication* (pp. 1–26). Oxford, UK: Oxford University Press.
- Harper, G., & Chitty, C. (2005). *The impact of corrections on re-offending: A review of "What Works"*. London: Home Office.
- Harre, R. (1972). *The philosophies of science: An introductory survey*. Oxford, UK: Oxford University Press.
- Hillage, J., Pearson, R., Anderson, A., & Tamkin, P. (1998). *Excellence in research on schools*. London: Department for Education and Employment.
- Hughes, J. (2004). *Doing the arts justice: A review of research literature, practice and theory*. The Unit for the Arts and Offenders, Centre for Applied Theatre Research. Retrieved April 1, 2009, from <http://www.apcentre.org.uk>
- Jackson, M.-R. (1998). Arts and culture indicators in community building: Project update. *Journal of Arts Management, Law and Society*, 28(3), 201–203.
- Jackson, M.-R., Herranz, J.Jr., & Kabwasa-Green, F. (2003). *Art and culture in communities: A framework for measurement* (Policy Brief No. 1 of the Culture, Creativity, and Communities Program). Washington, DC: The Urban Institute.
- Jackson, M.-R., Kabwasa-Green, F., & Herranz, J. (2006). *Cultural vitality in communities: Interpretation and indicators*. Washington, DC: The Urban Institute.
- Jermyn, H. (2001). *The arts and social exclusion: A review prepared for the Arts Council of England*. London: Arts Council England.
- Jermyn, H. (2004). *The art of inclusion* (Research Report No. 35). London: Arts Council England.
- Judge, K., & Mackenzie, M. (2002). Theory based evaluation: New approaches to evaluating complex community-based initiatives. In J. Mackenbach & M. Bakker (Eds.), *Reducing inequalities in health: A European perspective* (pp. 300–312). London: Routledge.
- Kilroy, A., Garner, C., Parkinson, C., Kagan, C., & Senior, P. (2007a). *Invest to save: Arts in health evaluation. Exploring the impact of creativity, culture and the arts on health and well being*. Manchester, UK: Arts for Health at Manchester Metropolitan University.
- Kilroy, A., Garner, C., Parkinson, C., Kagan, C., & Senior, P. (2007b). *Towards transformation: Exploring the impact of culture, creativity and the arts on health and wellbeing. A consultation report for the critical friends event September 2007*. Manchester, UK: Arts for Health at Manchester Metropolitan University.

- Landry, C., Bianchini, F., Henry, I., & Brookes, F. (2004). *Culture and regeneration: An evaluation of the evidence*. Culture East Midlands. Retrieved April 1, 2009, from <http://www.culture-em.org.uk>
- Landry, C., Bianchini, F., Maguire, M., & Worpole, K. (1993). *The social impact of the arts: A discussion document*. Stroud, UK: Comedia.
- Mackenzie, M., & Benzeval, M. (2005). Evaluating policy and practice: Designing the national HAZ evaluation. In M. Barnes, L. Bauld, M. Benzeval, K. Judge, M. Mackenzie, & H. Sullivan (Eds.), *Health Action Zones: Partnerships for health equity* (pp. 43–66). Abingdon, UK: Routledge.
- Macnaughton, J., White, M., & Stacy, R. (2005). Researching the benefits of arts in health. *Health Education*, 105, 332–339.
- Martin, S., & Sanderson, I. (1999). Evaluating public policy experiments: Measuring outcomes; monitoring processes or managing pilots? *Evaluation*, 5(3), 245–258.
- Matarasso, F. (1996). *Defining values: Evaluation in the arts*. Stroud, UK: Comedia.
- Matarasso, F. (1997). *Use or ornament? The social impact of participation in the arts*. Stroud, UK: Comedia.
- May, C. (2006). Health technology assessment and the politics of evidence. *Sociology of Health and Illness*, 28, 513–532.
- McDonnell, B., & Shellard, D. (2006). *Social impact of study of UK theatre*. London: Arts Council England.
- Merli, P. (2002). Evaluating the social impact of participation in arts activities: A critical review of Francois Matarasso's. *Use or Ornament? International Journal of Cultural Policy*, 8, 107–118.
- Michalos, A.C. (2005). Arts and the quality of life: An exploratory study. *Social Indicators Research*, 71, 11–59.
- Miell, D., MacDonald, R.A.R., & Hargreaves, D.J. (2005). *Musical communication*. Oxford, UK: Oxford University Press.
- Miles, A., & Clarke, R. (2006). *The arts in criminal justice: A study of research feasibility*. Manchester, UK: Centre for Research on Socio Economic Change, University of Manchester. Retrieved August 7, 2007, from http://www.culture.gov.uk/reference_library/publications/3559.aspx
- Millward, L., Kelly, M., & Nutbeam, D. (2003). *Public health intervention research: The evidence*. London: Health Development Agency.
- Moriarty, G. (1997). *Taliruni's travellers: An arts worker's view of evaluation – The social impact of the arts* (Working Paper No. 7). Stroud, UK: Comedia.
- Newman, T., Curtis, K., & Stephens, J. (2003). Do community-based arts projects result in social gains? A review of the literature. *Community Development Journal*, 38, 310–322.
- Oakley, K. (2004). *Developing the evidence base for support of cultural and creative activities in South East England*. Guildford, UK: South East England Cultural Consortium.
- Pawson, R. (2006). *Evidence-based policy: A realist perspective*. London: Sage.
- Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2004). *Realist synthesis: An introduction* (Methods Paper for the ESRC Research Methods Programme). ESRC Research Methods Programme. Retrieved April 1, 2009, from <http://www.ccsr.ac.uk>
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: Sage.
- Petticrew, M., & Roberts, H. (2005). *Systematic reviews in the social sciences*. Oxford, UK: Blackwell.
- Ragin, C.C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley, CA: University of California Press.
- Reeves, M. (2002). *Measuring the economic and social impact of the arts: A review*. London: Arts Council England.
- Rogers, P., Petrosino, A., Huebner, T.A., & Hacsí, T.A. (2000). Program theory evaluation: Practice, promise, and problems. *New Directions for Evaluation*, 87, 5–13.
- Ruiz, J. (2004). *A literature review of the evidence base for culture, the arts and sport policy*. Edinburgh, UK: Scottish Executive.
- Rychetnik, L., Frommer, M., Hawe, P., & Shiell, A. (2002). Criteria for evaluating evidence on public health interventions. *Journal of Epidemiology and Community Health*, 56, 119–127.
- Sanderson, I. (2000). Evaluation in complex policy systems. *Evaluation*, 6, 433–454.
- Sanderson, I. (2002). Making sense of “What Works”: Evidence based policy making as instrumental rationality? *Public Policy and Administration*, 17, 61–75.
- Sayer, A. (1984). *Method in social science: A realist approach*. London: Routledge.
- Selwood, S. (2006). A part to play? *International Journal of Cultural Policy*, 12, 35–53.
- Shaw, P. (1999). *The arts and neighbourhood renewal: A literature review to inform the work of the Policy Action Team 10*. London: DCMS.
- Shipman, M. (1997). *The limitations of social research* (4th ed.). Harlow, UK: Addison Wesley Longman.

- Spandler, H., Secker, J., Kent, L., Hacking, S., & Shenton, J. (2007). Catching life: The contribution of arts initiatives to recovery approaches in mental health. *Journal of Psychiatric and Mental Health Nursing*, 14, 791–799.
- Spencer, L., Ritchie, J., Lewis, J., & Dillon, L. (2003). *Quality in qualitative evaluation: A framework for assessing research evidence. A quality framework*. London: Government Chief Social Researcher's Office.
- Staricoff, R.L. (2004). *Arts in health: A review of the medical literature* (Research Report No. 36). London: Arts Council England.
- Thomas, J., Harden, A., Oakley, A., Oliver, S., Sutcliffe, K., Rees, R., et al. (2004). Integrating qualitative research with trials in systematic reviews. *British Medical Journal*, 328, 1010–1012.
- Thomson, H., Atkinson, R., Petticrew, M., & Kearns, A. (2006). Do urban regeneration programmes improve public health and reduce health inequalities? A synthesis of the evidence from UK policy and practice (1980–2004). *Journal of Epidemiology and Community Health*, 60, 108–115.
- Tooley, J., & Darby, D. (1998). *Education research: A critique. A survey of published educational research*. London: Office for Standards in Education.
- Weiss, C.H. (1972). *Evaluation research*. Englewood Cliffs, NJ: Prentice-Hall.
- Weiss, C.H. (1997). Theory-based evaluation: Past, present and future. *New Directions for Evaluation*, 76, 41–55.
- Weiss, C.H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19(1), 21–33.
- Weiss, C.H. (2000). Which links in which theories shall we evaluate? *New Directions for Evaluation*, 87, 35–45.
- White, M. (2004). Arts in mental health for social inclusion: Towards a framework for programme evaluation. In J. Cowling (Ed.), *For art's sake? Society and the arts in the 21st century* (pp. 75–99). London: Institute for Public Policy Research.
- White, M., & Angus, J. (2003). *Arts and mental health literature review: Addressing the evidence base from participation in arts and cultural activities*. Durham, UK: Centre for Arts and Humanities in Health and Medicine (CAHHM).
- White, T.R., & Rentschler, R. (2006, August). *Towards a new understanding of the social impact of the arts*, Paper presented at the international conference on the Arts in Society, Edinburgh, UK.
- Winner, E., & Cooper, M. (2000). Mute those claims: No evidence (yet) for a causal link between arts study and academic achievement. *Journal of Aesthetic Education*, 34, 11–75.
- Winner, E., & Hetland, L. (2000). The arts in education: Evaluating the evidence for a causal link. *Journal of Aesthetic Education*, 34, 3–10.
- Wykes, T. (2004). What kind of mental health research should we be doing in the 21st century? Mental health research networks. *Journal of Mental Health*, 13, 5–9.

Copyright of Cultural Trends is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.