

Journal of Music, Health, and Wellbeing

Journal Homepage: www.musichealthandwellbeing.co.uk



The bagatelle of human flourishing: Using realist evaluation to disentangle the multiple wellbeing benefits of participatory music activity

Andrew Fletcher¹ and Simon Hackett²

¹Durham University

²Newcastle University

Article Info.

Date Submitted:
January 2020

Date Accepted:
January 2021

Date Published:
March 2021

Abstract

The benefits of music participation are seen across health, education and social care contexts, and can help different people with different needs in different ways. As the boundaries between these contexts increasingly overlap, the need for flexible and effective evaluation has never been greater. Integrated health and social care, combined with an increasing requirement to demonstrate value for money, points to a need for more refined measurement tools that can be adapted to multiple contexts. This article explores realist evaluation as a means for generating better evidence to support participatory music interventions. It is illustrated using an evaluation of music-based programmes designed to increase wellbeing for adults and children with mental health issues and/or a learning disability. We draw on this 'proof of concept' to propose that realist evaluation offers a suitably flexible approach for disentangling the range of causal mechanisms at work in wellbeing interventions and can strengthen and consolidate the evidence for arts-based interventions.

Keywords: Music, wellbeing, realist evaluation, mental health, learning disability, complex needs, social interventions, arts and health

1. Introduction

The last decade has seen an increasing focus on wellbeing in the UK, contributed to in part by the OECD's 'Better Life Index' (2011) and an emerging austerity-linked mental health 'crisis' (Stewart, 2019). This has given rise to successive high-level (Prime Ministerial) initiatives such as Cameron's 'Happiness Agenda' (2010) and May's 'loneliness strategy' (2018). The trend is also set against the backdrop of an integrated care agenda (part of the NHS Five Year Forward View (2014)), which acknowledges the social determinants of health and so aims to take a more holistic (and efficient) approach to health and social care. One recent innovation that aims to enhance integration is 'social prescribing', where health professionals can refer patients to non-medical sources of support in their community such as gym membership, interest groups or participation in arts activities (Drinkwater et al. 2019). However, there is a call for more robust evaluation of the benefits of social prescribing (Bickerdike et al. 2017). The need for better evaluation of social programmes is also partly driven by the UK government's programme of financial austerity, initiated in 2010, which reduced spending on a wide range of public services (Wells, 2018). These cuts have been especially severe on complex social programmes, including arts for health programmes, which must produce clear evidence of their effectiveness or risk being decommissioned (All-Party Parliamentary Group on Arts Health and Wellbeing, 2017: 22).

Correspondingly, there has been renewed interest in the role of the arts in improving wellbeing outcomes. Groups such as the All-Party Parliamentary Group on Arts Health and Wellbeing are beginning to have an impact at the local level (for example, the North Tyneside Health and Wellbeing Board now bases its strategy on mapping local cultural activity against GP loneliness data). Public Health Englandⁱ has been paying increased attention to the evaluation of arts for wellbeing programmes (Daykin and Joss, 2016), while the World Health Organisation (WHO) recently produced a scoping review of evidence from over 3,000 studies on the role of the arts in improving health and wellbeing (Fancourt and Finn, 2019).

These combined factors raise the question: how do we produce robust evidence of the multiple complex benefits of arts participation in different social contexts? Public Health England highlighted that 'there are no clearly established evaluation frameworks for arts in health and wellbeing' (Daykin and Joss, 2016: 5). Given that arts-based interventions can have specifically therapeutic, pedagogic, or social aims, and that these contexts increasingly overlap, there is a need for evaluation approaches that enable insight into the complexities of an activity, how it is perceived by individuals, and the setting in which it occurs. Of these, theory-driven approaches (approaches that pursue the *unseen*, such as causal mechanisms) can offer potentially valuable insights (Daykin et al., 2017). One such approach, 'theory of change' (ToC) uses a 'logic map' approach to identify the *mechanisms* of a programme's success by examining how its inputs become outcomes (Weiss, 1995). A similar approach, 'Realist Evaluation' (RE), uses abductive reasoning (shifting between empirical data and abstract theory to infer the best available explanation) to develop and refine theories about how programmes and their underpinning mechanisms work. Refining ideas about these mechanisms enables an evaluator to express 'ever more detailed answers to the question of *why* a programme works, for *whom* and in *what* circumstances' (Pawson and Tilley, 1997: xvi). RE provides a way of approaching complex social programmes, and has gained traction in public health and social research (Marchal et al., 2012; Manzano, 2016) due to its accounting for context and ability to organise theories from an 'overabundance of explanatory possibilities' (Pawson, 2006: 19). Because 'social programmes are always complex systems thrust amidst complex systems' (ibid.: 25), statutory bodies and government departments, such as NIHR (Harris et al., 2015) and HM Treasury (2020), are increasingly using or commissioning RE-type studies to address specific social policy research needs. However, evidence-based policymaking still prioritises systematic reviews and randomised controlled trials despite the fact that these are often inappropriate for evaluating complex social interventions (Cartwright and Hardie, 2012). Instead, RE offers a more nuanced perspective, which this article promotes as an appropriate method for evaluating music's impact on wellbeing.

The connections between music and all forms of human flourishing and wellbeing are well-documented (Ockelford, 2008; MacDonald et al., 2013; Bartleet and Higgins, 2018) and include a wide range of more specific benefits, from increasing concentration (Huang and Shih, 2011), reducing criminality (Tuastad and O'Grady, 2013), improving learning (Green, 2009), reducing somatic pain (Kühlmann et al., 2018), improving community cohesion (Veblen, 2008), aiding recall (Wallace, 1994), calming anxiety (Lee et al., 2012), providing relief from neurological conditions (Sacks, 1973) and many others. In most social contexts, music-based approaches can yield multiple positive outcomes through multiple linked mechanisms. As Byrne (2011: 29) notes: 'In our discussion of a realist version of causality, we should consider that there are multiple mechanisms which can generate the same outcome'. In such complex systems, it is not always clear *which* mechanisms are active in which contexts, and which outcomes are important to individuals. Approaches that attempt to (at least partially) make sense of this tangle can therefore provide a useful focus. RE provides a framework to explore these and to marshal explanations of: 'What works, for whom, in what respects, to what extent, in what contexts, and how?' (Pawson and Tilley, 1997: xvi). It allows the consideration of multiple theories simultaneously and, though an iterative process, enables these to be refined, developed and disentangled in relation to their context and each other.

Drawing on a study by Fletcher (2017), which examined the range of ways music participation could increase wellbeing for people in challenging circumstances, this article explores the use of RE as an important method for unpicking the mechanisms at play across intersecting contexts. In recognising that the benefits of music are leveraged across *and between* contexts, we consider RE's flexibility in this type of trans/multi-contextual work. At the time of writing, RE has few references among studies of arts and health (Sridharan and Nakaima, 2012; Froggett and Roy, 2014; Madsen, 2018) and fewer still in music and wellbeing (Fletcher, 2017; Porter et al., 2017a; Caló et al., 2020). We propose that such studies can be seen as a 'proof of concept' for using RE in arts and health. We also consider some objections to the methodology and caution against seeing it as panacea to the problem of evaluating arts-based interventions. As a means to gain deeper understandings of mechanisms, with a view to carrying out interventions more effectively and efficiently, RE merits greater attention from those seeking to develop the evidence base for music initiatives and interventions across sectors.

2. Realist Evaluation - overview

Realist Evaluation is a theory-driven evaluation framework developed by Pawson and Tilley (1997). Emerging from the realist traditions in sociology (Outhwaite, 1987; Sayer, 1992; Bhaskar, 2013), RE identifies, develops, tests and refines programme theories – discrete hypotheses about what is happening in an intervention, or 'units of explanatory potential' (Fletcher, 2017). Initial programme theories are derived from a range of sources including academic literature, policy documents, anticipated programme outputs, observation and elicitation. These are then supported and developed, or refuted, through the iterative accrual of data. Some programme theories might emerge more strongly than anticipated, meriting further investigation, while others might be discarded because they are not triggered within that programme. This does not render the latter invalid; they just weren't actualised in the particular context being studied. In a different context, they might be observed and pursued. Similarly, unexpected outcomes might lead to new programme theories being developed and explored within the evaluation.

RE accepts both positivist and interpretivist evidence (treating formalised programme outputs and participants' experiences equally), allowing for the development of programme theories that shed greater light on the complex systems inherent within social programmes. By permitting a range of data and data collection techniques, RE is potentially more accessible to non-specialist evaluators. 'The data used to develop, and test explanations can be either quantitative or qualitative. Realist evaluators are generally agnostic with respect to types of data' (Mathison, 2005: 361). Fletcher (2017) used primarily qualitative methods, with the aim of identifying and exploring generative mechanisms that connect music

activities with increased wellbeing. Different studies might use different data collection methods according to their aims (Daykin et al., 2017).

Programme theories are fluid models describing fluid situations. They are refined by combining multiple data ‘snapshots’ and iteratively amending the theory to develop a sharper image (abductive reasoning). The success of any intervention in a social context depends on the extent to which the programme theory/ies predicted or controlled the spiral of ideas and changes that occurred because of that intervention. Predictability indicates a level of consistency and therefore a successful intervention design. Therefore, enough data snapshots must be taken so that the programme theory can describe ‘demi-regularities’, or ‘semi-predictable patterns or pathways of programme functioning’ (Dieleman et al., 2012: 27). Demi-regularities strengthen the likelihood of a programme theory being applicable in a different intervention, although it always remains open to further refinement.

Programme theories can be articulated as Context Mechanism Outcome Configurations (CMOCs), based on the ‘trio of explanatory components’ (Pawson and Tilley, 1997: 77) that marshal theories into relatively standardised units. This heuristic device seeks to address the consistency problem that has historically dogged social sciences. Using CMOCs to articulate programme theories more consistently makes it easier to test and explore them in different contexts. CMOCs therefore increase the ‘portability’ of theories, enabling deeper understandings of the links between context and outcome. Social programmes are considered successful (in their ‘outcomes’) if they introduce the appropriate ideas and opportunities to groups in the appropriate social and cultural conditions (‘contexts’), thereby triggering a ‘mechanism’. It is therefore reasonable to ask: will *these* contextual resources trigger *this* mechanism to yield *this* outcome? RE flows from this explanatory preposition; identifying and developing theories, testing, and refining these through the intervention, then producing CMOCs to further interrogate the data or be tested in different contexts.

Fletcher (2017) asked: outside of specific music therapy settings, how do we know *which* aspects of participatory music’s impacts on wellbeing are active in which contexts? This is a recognised issue; the 2015 National Institute for Clinical Excellence (NICE) guidelines on older people and mental wellbeing state: ‘The committee discussed the evidence on singing and noted that it is unclear whether it is the singing itself that produces the benefit, the group-based nature of the activity or something else’ (NICE, 2015: 23). It is exactly this type of question that RE is designed to address, to produce reliable and transferrable evidence that can be used and developed by different organisations across sectors.

3. Illustrative study of a realist evaluation of music for wellbeing

Fletcher (2017) looked at three music-based interventions at three study sites: 1) An NHS inpatient unit for children aged 12 to 18 with complex mental health needs and/or a learning disability. The unit provided a weekly optional group music programme. Song writing activities, a recording project, drumming workshops and opportunities for live performance were delivered by a music therapist and a community musician. 2) A charitable organisation providing a range of arts-based resources for adults with (or recovering from) mental health issues. This included a music room containing musical instruments and recording equipment. Bi-weekly facilitated sessions involved jamming, song writing and personal recording projects, chosen by the participants. 3) An educational group, training young adults (aged 16 - 18) with learning disabilities to become community music leaders. The group devised new musical activities, recorded and produced original material, and staged public concerts and outreach events. All participants faced various challenges and had chosen music creation as a way of increasing their wellbeing. The inpatient site was the most controlled environment, and activities were carefully planned and designed by the facilitators; the charitable organisation’s music activities were mostly participant led, while the educational group used democratic decision-making within a curricular framework.

All participants agreed that their involvement in music activities was beneficial to their wellbeing but each gave different reasons. This ‘overabundance of explanatory possibilities’ led to a long list of ‘candidate’ programme theories – 31 in total – based on a wide, shallow literature review. After participating in each of the three groups for several weeks and talking informally to group members, the researcher was able to narrow the list. For example, there is a literature on the links between music and spirituality, but it emerged that spirituality was not a factor for these participants, so that programme theory was excluded. This stage was followed by semi-structured ‘realist interviews’, which are explicitly theory-driven and investigate propositions about how, where, when and why programmes are and are not effective (Manzano, 2016; Mukumbang et al., 2020). The shortlist of programme theories was explored with a verification group (drawn from the third case study) to establish if these were accurate representations of participants’ experiences and to gain further comments or perspectives. This data was combined with further literature until several themes were able to be refined and formulated as CMOCs (summarised in Table 1). While the broad programme theories already have a place in existing literature, examining how these played out and interacted in different contexts gave a useful insight into how interventions might be tailored to their participants or circumstances.

Table 1. Simplified CMOCs from Fletcher (2017)

Broad theory	Observed Context Mechanism Outcome configuration
Praise and hope	Group music activity (context) enabled opportunities for participants to receive praise. At the inpatient unit (context), <u>praise from adults</u> gave rise to a sense of hope for recovery (mechanism); at the charity (context), <u>approbation from peers</u> (mechanism) was deemed to increase wellbeing (outcome) more directly.
Energy control	Improvised music activity could increase or decrease perceived energy levels, giving participants a sense of control they otherwise felt they lacked.
Representation of self	Recording their music allowed participants to represent themselves to the wider world. Some used this to connect with family and friends; others to assert their personhood in a more political way.
Genre and subculture	Having choices allowed participants to align themselves with – or against – a mainstream style, enabling them to position their identity in relation to the wider world. As above, this brought a sense of ‘having a voice’.
Resilience	Effects of music participation lasted beyond the end of the activity. These effects were interpreted in different ways but many participants deemed this to increase their ‘resilience’ (for example, giving them confidence to use public transport).
Memory	Listening to and participating in music triggered specific memories for some participants or evoked emotions. Combined with musical choices, this led to positive feelings, increasing wellbeing.

These findings represent the combination of wellbeing mechanisms that were being triggered for those particular musicians at those particular study sites. They: 1) provide immediate information about wellbeing mechanisms and what types of music activity or other contextual factors enabled these, helping to focus existing work; and 2) give a good indication about what might work in similar contexts, providing a starting point for designing or refining other music programmes. This is important; understanding what might work in similar contexts, and why, is critical to developing efficient and effective future arts and health interventions that can be funded with confidence. This evaluation led to the development of a model for the design and facilitation of similar music activities (Fletcher et al., 2019), which informs current arts programming at the inpatient unit, and has led to two successful funding bids to the National Foundation for Youth Music.

We assert that, although this study had a relatively small sample of self-selected (for example, individuals who were already attracted to music activity) participants, it nevertheless illustrates the value of RE in teasing out specific benefits of art interventions.

4. Discussion

As an increasingly popular methodological framework in social research, there are ongoing debates around RE. Many of these are explored within the Realist And Meta-narrative Evidence Syntheses: Evolving Standards (RAMESES) projects, whose website and mailing lists are a popular resource for discussing realist theory and study designs.ⁱⁱ Here, we explore some of the main issues.

Transferability

The transferability of findings is vital to ensure that research carried out in one context can have a beneficial impact in similar contexts. In our example, how do we know that a wellbeing mechanism triggered by participatory music activity in one place will be triggered in the same way elsewhere? While generalisability remains an intractable problem in social science, especially given the range of data types and collection methods available, the CMOC heuristic enables a certain degree of standardisation and so portability of findings.

Data collection is usually governed by the type of questions being asked. Evaluating the dappled realm of subjective wellbeing, with its different understandings and constructions, achieved through a ‘bagatelle’ of context-dependent pathways, often requires a mixed-methods approach. In our illustrative study, quantitative data about context (group size; length of sessions; age ranges) and outcomes (wellbeing scores based on the Warwick-Edinburgh Mental Wellbeing Scale) was combined with qualitative inquiry around mechanisms (exploring candidate programme theories as well as participants’ own reasoning as to why the intervention was beneficial). CMOCs are developed out of a ‘conversation’ between the available data and an iteratively refined programme theory. The products of the enquiry, CMOCs enable specific findings and programme theories to be articulated in a way that has both ontological depth (consistent with Bhaskar’s ‘stratified’ model (2013: 13)) and pragmatic reliability, in the sense of being portable.

Because programmes work differently in different contexts and through different change mechanisms, we cannot assume that programmes can be replicated from one context to another or that they will automatically achieve the same outcomes if they are. What is portable, however, are good understandings about ‘what works for whom, in what contexts, and how’ (Westhorp, 2014: 7 - quoting Pawson and Tilley, 1997).

This portability is also made possible because CMOCs tend to represent ‘middle-range theories’ – they are neither too narrow (only applying in very specific contexts) nor too general (providing overarching ‘grand’ theories). Pawson and Tilley deem middle-range theories to ‘...provide analytic frameworks to interpret similarities and differences between families of programmes’ (1997: 217), indicating the level of similarity at which contexts might be usefully compared. In our example, that means organised group music activities designed to bring about wellbeing-related changes; these are more general than focused music therapy interventions and more specific than everyday musical encounters.

Theory refinement

Abductive reasoning, which underpins RE, does not claim absolute truth, merely ‘inference to the best explanation’ (Sober, 2001: 28). Individual theories are open to ongoing refinement. Further to this, multiple programme theories might be active to greater or lesser extents, depending on the context. For example, the music programme at the inpatient setting in our example study gave rise to therapeutic, educational and social outcomes, enabling two fronts for progress:

1) In terms of individual programme theory refinement, we return to CMOCs. Contexts, mechanisms and outcomes exist in a dynamic *configuration*; if one changes, so do the others. This fluidity means that outcomes are not necessarily replicable in exactly the same way in different contexts. Instead, ‘Rather than replicate interventions in anticipation of the same results, the realist evaluator sees subsequent trials as an opportunity for CMO *configuration focusing*’ (Pawson and Tilley, 1997: 217 – italics in original). This is the essence of programme theory refinement; small changes in C, M or O revealed by the data give rise configuration changes in the programme theory, which evolves in relation to the programme it is tested in. Consequently, RE does not seek outcome *regularities* (replicability) across contexts but instead examines outcome patterns in order to develop a more comprehensive understanding of the generative causal mechanism (Salter and Kothari, 2014).

Evaluators must choose between two epistemological standpoints: Critical Realism assumes ‘an overabundance of explanatory possibilities’ (Pawson, 2006: 19), leading the researcher to be persistently critical of potentially false or mistaken explanations (Bhaskar, 2013) and to pursue ever more refined theories. Conversely, Empirical Realism acknowledges the need to ‘draw a line’ and select explanations in spite of the fact that further potential explanatory mechanisms may be uncovered (Pawson, 2006) – what Bhaskar calls the TINA (there is no alternative) aspect (2008: 116). The latter is a more pragmatic approach, especially if evidence is to be *usable* by both policymakers and funders.

2) When multiple programme theories are active, these may emerge in different proportions or interrelate differently with one another in different contexts. Our illustrative example was further developed into four themes that were evident to greater or lesser degrees across all three music groups: ‘peers’, ‘personhood’, ‘product’ and ‘positive interaction’. These formed the basis for the *4P participatory arts recovery model* (Fletcher et al., 2019), which encourages music programme designers and facilitators to focus activities in a way that reflects the particular needs of the group. The *4P* model draws from a ‘stock’ of CMOCs developed through empirical research but its configuration differs as the context changes (adults/children; therapeutic/educational; varying social dynamics etc.).

Numerous theories, set within a broad literature, connect arts participation with desired outcomes. RE of more real-life arts interventions would generate a stock of ‘off the peg’ yet flexible programme theories that can be combined, refined or discarded. While ongoing theory refinement might in some respects be a specialist or research job, having a stock of flexible and commonly observed CMOCs that can be used in a range of contexts would help practitioners and policymakers navigate and organise the dense evidence landscape.

Intersecting disciplines

The specific benefits of arts participation have been explored from within a range of disciplines, from experimental psychology to ethnographic community studies. However, there is less understanding of how more discipline-specific (neurological, psychological, social etc.) mechanisms work in cohesion, and of their concurrent and interacting nature. While edited volumes have noted the ‘...lack of coherence of research in this field. The diversity of approaches and findings, the heterogeneity of methods, participants, outcomes and interpretations of findings...’ (MacDonald et al., 2013: 4), these are often only marshalled into a navigable book structure; few develop a theory of how the findings from different approaches fit together. Disciplinary siloization might have once been implicated. For example, psychodynamic music therapy and community music therapy may yield similar benefits for participants via different mechanisms, active to greater or lesser degrees. But while those two approaches are not incompatible, they would tend to explain their effects differently. In a research milieu increasingly characterised by complexity-based and interdisciplinary approaches, methodologies that can account and combine multiple explanatory mechanisms have a great deal to offer.

Criticisms of and debates within realist evaluation

Much of the criticism of RE focuses on the framework's philosophical underpinnings. Sam Porter, a health sociologist with an interest in arts-based therapies, draws attention to RE's propensity to confuse mechanism and context, and its subsuming of agency under structure (2015). These types of theoretical inconsistency are reflected in our illustrative study, which did not always make clear the distinction between the general causal mechanisms on which the programme was based and why the *participants* believed the intervention was beneficial. For example, where recording apparatus (a resource) was provided to participants, wellbeing outcomes were observed as expected but participants' reasoning as to why this was differed according to context: older participants (especially those who had experienced problems socialising) valued the ability to represent themselves through a musical product; younger participants found joy in the novelty of the recording process; others were unaware of the causal mechanisms that led to their increased wellbeing. These ideas were explored in the interviews but untangling context from mechanism was not always possible. Dalkin et al. (2015: 3) describe Pawson and Tilly's original conceptualisation of mechanisms as 'a combination of resources offered by the social programme under study and stakeholders' reasoning in response'. They develop the theory further by disaggregating resources and reasoning – but retaining their causal relationship – within mechanisms. This helps to tease out and draw a firmer distinction between context and mechanism (Dalkin et al., 2015), making programme theories more refined, sensitive and useful.

The broader issue here is about how RE can hold both empiricist and idealist positions. Elder-Vass (2012) argues that it is possible for critical realists reconcile and maintain coherent versions of both realist and social constructionist ontologies simultaneously, while Porter raises the possibility of critical realist randomised controlled trials (but maintains that RE approaches are incompatible with RCTs) (Porter et al., 2017b). This at least reflects efforts to push the debate forward and move realist approaches up the social policy evidence hierarchy.

While concerns about mechanisms/contexts, facts/values and realism/idealism are worthy of debate, they do not necessarily diminish the useful products of RE. Arguably, RE works well at a small-scale programme level but its unresolved theoretical issues become more problematic at larger scale or organisation levels. Pawson believes these types of high-level objections are overly philosophical and abstract (2016), and the argument over RE's theoretical robustness versus its pragmatic use rumbles on (Porter, 2017; Porter et al., 2017b). Others have dismissed the argument as merely 'the continuation of a sociologists' turf war' (Hind, 2016: 1). We take an 'empirical realist' position; the findings from Fletcher (2017) coalesced into a flexible model, which has been adopted and used in similar contexts but remains open to further development. According to Blaikie (2007: 25), 'For [conventionalists], the truth status of the theories used to understand and manipulate the world of objects and events is not important. Rather, it is what such theories allow us to do that matters'. We share this conventionalist perspective. Concerns about inconsistent evidence and consequent lack of funding for the arts (despite high interest) are essentially pragmatic. In a policy and research environment where 'impact' has dominance, the ability to develop easy to implement and portable models that improve through sharing and collating empirical data, is both apposite and valuable.

5. Conclusion

Understanding the relationship between arts activity and wellbeing is nuanced and complex, as is the evaluation of social programmes more generally. RE is underused in this area but provides a valuable theoretical framework that can help elucidate understandings of the benefits of arts participation across a range of health, social care and educational settings. Its products, in the form of Context Mechanism Outcome Configurations, have the potential for more practical applications in programme design, while also producing useful and accessible evidence for arts and culture policymakers.

Given the increasing demands on arts programmes to generate evidence and demonstrate outcomes, including cost-effectiveness, we propose RE as a quick and easy way to fulfil these aims and increase understanding generally. The example study described here illustrates that the framework can be applied but we also recognise the need for further evaluations of this type to explore, refine and organise the evidence. This would 1) enable more efficient and effective use of resources in music and arts-related social programmes and 2) help generate more convincing arguments to persuade policymakers of the value of music and the arts to increase human flourishing across contexts.

References

- All-Party Parliamentary Group on Arts Health and Wellbeing 2017. *Creative health: the arts for health and wellbeing*. [Available online: <https://www.artshealthresources.org.uk/docs/creative-health-the-arts-for-health-and-wellbeing>] Last Accessed: 1 December 2020.
- Bartleet, B. L. and Higgins, L. 2018. *The Oxford handbook of community music*. New York: Oxford University Press.
- Bhaskar, R. 2008. *Dialectic: The pulse of freedom*. Abingdon: Routledge.
- Bhaskar, R. 2013. *A realist theory of science*. Abingdon: Routledge.
- Bickerdike, L., Booth, A., Wilson, P.M., et al. 2017. 'Social prescribing: less rhetoric and more reality. A systematic review of the evidence'. *BMJ open* 7(4). [Available online: <https://bmjopen.bmj.com/content/7/4/e013384>].
- Blaikie, N. 2007. *Approaches to social enquiry: Advancing knowledge*. Cambridge: Polity.
- Byrne, D. 2011. *Applying social science: The role of social research in politics, policy and practice*. Bristol: Policy Press.
- Caló, F., Steiner, A., Millar, S., et al. 2020. 'The impact of a community-based music intervention on the health and well-being of young people: A realist evaluation'. *Health & Social Care in the Community*. 28(3) pp.988-997.
- Cartwright, N. and Hardie, J. 2012. *Evidence-based policy: A practical guide to doing it better*. Oxford: Oxford University Press.
- Dalkin, S.M., Greenhalgh, J., Jones, D., et al. 2015. 'What's in a mechanism? Development of a key concept in realist evaluation'. *Implementation science*. 10(1) pp.1-7.
- Daykin, N., Gray, K., McCree, M., et al. 2017. 'Creative and credible evaluation for arts, health and well-being: opportunities and challenges of co-production'. *Arts & Health*. 9(2) pp.123-138.
- Daykin, N. and Joss, T. 2016. *Arts for health and wellbeing: An evaluation framework*. London: Public Health England.
- Dieleman, M., Wong, G. and Marchal, B. 2012. *An introduction to realist research*. Workshop presentation at the Second Global Symposium on health systems research, Beijing, 31 Oct - 3 Nov 2012. [Available online: https://www.abdn.ac.uk/femhealth/documents/Realist_methods_workshop.pdf] Last Accessed 1 December 2020.
- Drinkwater, C., Wildman, J. and Moffatt, S. 2019. 'Social prescribing'. *Bmj*. 364 pp.11285.
- Elder-Vass, D. 2012. *The reality of social construction*. Cambridge: Cambridge University Press.
- Fancourt, D. and Finn, S. 2019. *What is the evidence on the role of the arts in improving health and well-being? A scoping review*. [Available online: <https://europepmc.org/article/NBK/nbk553773>] Last Accessed: 10 December 2020.
- Fletcher, A. 2017. *A realist evaluation of participatory music interventions for wellbeing: What works, for whom and in what circumstances?* Doctoral Thesis, Northumbria University, Newcastle upon Tyne.

- Fletcher, A., Hackett, S. and Carr, S.M. 2019. 'The 4P participatory arts recovery model: Peers, product, personhood and positive interaction'. *Journal of Applied Arts & Health*. 10(1) pp.41-56.
- Froggett, L. and Roy, A. 2014. *Cultural attendance and public mental health: Evaluation of pilot programme 2012-2014*. [Available online: <http://clock.uclan.ac.uk/10962/1/CAPMH%20final%20report%20v91.pdf>] Last Accessed: 1 December 2020.
- Gov.uk 2010. *PM speech on wellbeing*. [Available online: <https://www.gov.uk/government/speeches/pm-speech-on-wellbeing>] Last Accessed: 1 January 2021.
- Gov.uk 2018. *PM launches Government's first loneliness strategy*. [Available online: <https://www.gov.uk/government/news/pm-launches-governments-first-loneliness-strategy>] Last Accessed: 1 January 2021.
- Green, L. 2009. *Music, informal learning and the school: A new classroom pedagogy*. Aldershot: Ashgate.
- Harris, J., Springett, J., Booth, A., et al. 2015. 'Can community-based peer support promote health literacy and reduce inequalities? A realist review'. *Public Health Research* 3(3). [Available online: <https://www.journalslibrary.nihr.ac.uk/phr/phr03030/#/abstract>].
- Hind, D. 2016. *Reviewer's Report: The Possibility of Critical Realist Randomised Controlled Trials*. [Available online: https://static-content.springer.com/openpeerreview/art%3A10.1186%2Fs13063-017-1855-1/13063_2017_1855_ReviewerReport_V0_R1.pdf]. Last Accessed: 1 December 2020.
- HM Treasury 2020. *Magenta Book 2020. Supplementary Guide: Realist Evaluation*. [Available online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/879435/Magenta_Book_supplementary_guide._Realist_Evaluation.pdf]. Last Accessed: 1 December 2020.
- Huang, R. H. and Shih, Y. N. 2011. 'Effects of background music on concentration of workers'. *Work*. 38(4) pp.383-387.
- Kühlmann, A., De Rooij, A., Kroese, L., et al. 2018. 'Meta-analysis evaluating music interventions for anxiety and pain in surgery'. *British Journal of Surgery*. 105(7) pp.773-783.
- Lee, K.C., Chao, Y.H., Yiin, J.J., et al. 2012. 'Evidence that music listening reduces preoperative patients' anxiety'. *Biological research for nursing*. 14(1) pp.78-84.
- MacDonald, R., Kreutz, G. and Mitchell, L. 2013. *Music, health, and wellbeing*. Oxford: Oxford University Press.
- Madsen, W. 2018. 'Raising social consciousness through verbatim theatre: a realist evaluation'. *Arts & Health*. 10(2) pp.181-194.
- Manzano, A. 2016. 'The craft of interviewing in realist evaluation'. *Evaluation*. 22(3) pp.342-360.
- Marchal, B., Van Belle, S., Van Olmen, J., et al. 2012. 'Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research'. *Evaluation*. 18(2) pp.192-212.
- Mathison, S. 2005. *Encyclopedia of evaluation*. Thousand Oaks, Ca: Sage.
- Mukumbang, F.C., Marchal, B., Van Belle, S., et al. 2020. 'Using the realist interview approach to maintain theoretical awareness in realist studies'. *Qualitative Research*. 20(4) pp.485-515.
- National Institute for Clinical Excellence 2015. *NICE Guideline [NG32] Older people: independence and mental wellbeing*. [Available online: <https://www.nice.org.uk/guidance/ng32>]. Last Accessed: 1 December 2020.
- NHS England 2014. *Five Year Forward View*. [Available online: <https://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf>]. Last Accessed: 1 December 2020.
- Ockelford, A. 2008. *Music for children and young people with complex needs*. Oxford: Oxford University Press.
- Organisation for Economic Co-Operation and Development. 2011. *Better Life Index*. [Available online: <http://www.oecdbetterlifeindex.org/>] Last Accessed: 1 January 2021.
- Outhwaite, W. 1987. *New philosophies of social science: realism, hermeneutics and critical theory*. Basingstoke: Macmillan Education Ltd.

- Pawson, R. 2006. *Evidence-based policy: a realist perspective*. London: Sage.
- Pawson, R. 2016. 'Realist evaluation caricatured: a reply to Porter'. *Nursing Philosophy*. 17(2) pp.132-139.
- Pawson, R. and Tilley, N. 1997. *Realistic Evaluation*. London: Sage.
- Porter, S. 2015. 'Realist evaluation: an immanent critique'. *Nursing Philosophy*. 16(4) pp.239-251.
- Porter, S. 2017. 'Evaluating realist evaluation: a response to Pawson's reply'. *Nursing Philosophy* 18(2) p.12155.
- Porter, S., McConnell, T., Clarke, M., et al. 2017a. 'A critical realist evaluation of a music therapy intervention in palliative care'. *BMC palliative care*. 16(1) p.70.
- Porter, S., McConnell, T. and Reid, J. 2017b. 'The possibility of critical realist randomised controlled trials'. *Trials*. 18(1) p.133.
- Sacks, O. 1973. *Awakenings*. London: Duckworth.
- Salter, K.L. and Kothari, A. 2014. 'Using realist evaluation to open the black box of knowledge translation: a state-of-the-art review'. *Implementation science*. 9(1) p.115.
- Sayer, A. 1992. *Method in social science: A realist approach*. London: Routledge.
- Sober, E. 2001. *Core questions in philosophy: A text with reading*. 3rd edn, London: Pearson.
- Sridharan, S. and Nakaima, A. 2012. 'Towards an evidence base of theory-driven evaluations: Some questions for proponents of theory-driven evaluation'. *Evaluation*. 18(3) pp.378-395.
- Stewart, M. 2019. 'Preventable harm: creating a mental health crisis'. *Journal of Public Mental Health*. 18(4), pp.224-230.
- Tuastad, L. and O'Grady, L. 2013. 'Music therapy inside and outside prison—A freedom practice?' *Nordic journal of music therapy*. 22(3) pp.210-232.
- Veblen, K.K. 2008. 'The many ways of community music'. *International Journal of Community Music* 1(1) pp.5-21.
- Wallace, W.T. 1994. 'Memory for music: Effect of melody on recall of text'. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 20(6) p.1471.
- Weiss, C.H. 1995. 'Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families'. *New approaches to evaluating community initiatives: Concepts, methods, and contexts*. 1 pp.65-92.
- Wells, P. 2018. 'Evidence based policy making in an age of austerity'. *People, place and policy online*. 11(3) pp.175-183.
- Westthorp, G. 2014. 'Realist impact evaluation: an introduction'. *London: Overseas Development Institute*. pp.1-12.

ⁱ Since this article was written, the Covid-19 pandemic has caused the UK government to announce that Public Health England will be replaced by the National Institute for Health Protection. It is currently unclear what will become of PHE projects relating to arts and health. The consequent lockdown, including the closure of live music venues, has also refocussed debates around the value of arts and music to wellbeing. However, the economic impact of the pandemic will undoubtedly lead to further spending cuts.

ⁱⁱ <https://www.ramesesproject.org/>