Creative and credible evaluation for arts, health and well-being: opportunities and challenges of co-production

Norma Daykin, Karen Gray, Mel McCree & Jane Willis

Published online: 26 Sep 2016.

Submit your article to this journal ⍙

Article views: 1295

View related articles ⍙

View Crossmark data ⍙

Taylor & Francis
Taylor & Francis Group

Creative and credible evaluation for arts, health and well-being: opportunities and challenges of co-production

Norma Daykin[a] , Karen Gray[b], Mel McCree[c] and Jane Willis[d]

[a]Centre for Arts as Wellbeing, University of Winchester, Winchester, UK; [b]TAnDem Doctoral Training Centre and the Association for Dementia Studies, University of Worcester, Worcester, UK; [c]Department of Health and Social Sciences, University of the West of England, Bristol, UK; [d]Willis Newson, Bristol, UK

ABSTRACT

Background: This paper reports findings from a one-year UK knowledge exchange (KE) project completed in 2015. Stakeholders' experiences of evaluation were explored in order to develop online resources to strengthen knowledge and capacity within the arts and health sector (www.creativeandcredible.co.uk).
Methods: The project used mixed methods, including a survey, interviews and focus groups, guided by a Stakeholder Reference Group comprised of 26 leading UK evaluators, researchers, artists, health professionals, commissioners and funders.
Results: The project identified opportunities for arts arising from current health and social care policy agendas. It also identified challenges including the lack of agreed evaluation frameworks and difficulties in evaluation practice.
Conclusions: Co-production between stakeholders is needed to strengthen evaluation practice and support the development of the arts and health sector. Effective co-production can be undermined by structural and cultural barriers as well as unequal stakeholder relationships. The paper discusses recent initiatives designed to support best practice.

Background

This paper reports findings from a UK knowledge exchange (KE) project that ran from September 2014 to August 2015 (Daykin, Willis, Gray, & McCree, 2016). The project grew out of a long-standing collaboration between a University and a leading arts and health consultancy focused on improving evaluation practice in arts, health and well-being sector.

Since the review of the international evidence by Rosalia Staricoff (2004), updated by Staricoff & Clift (2011), research in this field has continued to grow. A significant number of studies have focused on music and singing (Baird & Samson, 2015; Camic, Williams, & Meeten, 2013; Daykin, de Viggiani, Pilkington, & Moriarty, 2013; Osman, Tischler, & Schneider, 2014; Simpson & Keen, 2011; Skingley et al., 2014; Zhang et al., 2012). A growing body of research has also examined other art forms, including visual arts (Camic, Tischler, & Pearman, 2014; Cusack, Lankston, & Isles, 2010; Daykin, Byrne, Soteriou, & O'Connor, 2010; Gelo, Klassen, &

Gracely, 2015; Nanda et al., 2012; Sonke et al., 2015); performing arts (Daykin et al., 2008); creative writing (Gregory, 2011; Rickett, Greive, & Gordon, 2011); and museum object handling (Chatterjee & Camic, 2015; Johnson, Culverwell, Hulbert, Robertson, & Camic, 2015).

As well as hospital-based provision, there is growing awareness of the potential for arts in other areas including community services (Fraser, Bungay, & Munn-Giddings, 2014; Sonke & Lee, 2015), as well as prevention and public health (Clift, 2012; Kreutz & von Ossietzky, 2015; McDaid & Park, 2013; Renton et al., 2012). In the UK, where this study was based, devolved commissioning of services to meet locally determined needs following the 2012 Health and Social Care Act has created new opportunities for innovations across these areas, building on established programmes of arts on prescription, where health and social care practitioners can refer people to an arts activity or service as a source of support (Bungay & Clift, 2010) The Arts Council England's Cultural Commissioning programme (www.ncvo.org. uk/practical-support/public-services/cultural-commissioning-programme) has engaged the arts and cultural sector in local commissioning, further strengthening advocacy for arts in health and well-being.

While this policy context is specific to the UK, it has reinforced challenges that exist in many contexts where arts and health projects and programmes are in development, particularly in relation to programme evaluation and evidence generation. Hence as well as creating opportunities, recent developments have increased demands on projects to demonstrate outcomes, including cost-effectiveness.

Despite advances in research, development of the sector has been hindered by a lack of consensus about how to evidence the value of arts, particularly at the level of practice, leading to a perception that the evidence base is weak (Carnwath & Brown, 2014; Mowlah, Niblett, Blackburn, & Harris, 2014; Skingley, Bungay, & Clift, 2011). Artists often report feeling marginalized by evaluation discourse, which is dominated by methodologies drawn from health and social sciences (Daykin, Attwood, & Willis, 2013). The project sought to address this need by developing freely available online resources, available to arts organizations seeking to develop their evaluation practice, as well as funders, commissioners and policy-makers. These web-based resources are now available at www.creativeandcredible.co.uk. They are designed to enable stakeholders to broaden their evaluation knowledge and skills in order to effectively engage with policy and funding agendas. Here, we report on the process of research and development that led to the creation of the website, including underpinning research with leading UK arts and health organizations, researchers and evaluators.

Research design and methods

The research sought to explore current approaches and experiences of evaluation across the arts and health sector, identifying key challenges and stakeholder concerns in order to inform the development of appropriate resources. The project was guided by a Stakeholder Reference Group (SRG) comprised of 26 leading evaluators, researchers, artists and arts organizations, health professionals, and health and social care commissioners, who met in a series of five seminars and workshops. Reflections on the SRG discussions shaped data collection methods and informed data analysis.

Additional methods included interviews and a descriptive online survey using the Bristol Online Survey facility (www.onlinesurveys.ac.uk). This simple survey included 10 questions

designed to capture stakeholders' experiences of evaluation. Respondents were asked about their current roles in arts and health and their experience of using quantitative, qualitative and creative arts-based evaluation methods. They were asked to provide details of evaluation projects that they had undertaken over the last five years including topic, funding, duration, methodology and publication details. They were also asked about their experiences of working with commissioners and funders when devising and delivering project and programme evaluations. The survey was sent to a purposeful sample of 26 evaluation stakeholders identified through our SRG and we also placed an open invitation to take part on the project website.

Semi-structured interviews were undertaken over the phone or in person with nine stakeholders selected on advice from our SRG and drawn from service providers (2), researchers (2), evaluators (2) and commissioners (3). Interviews lasted between 20 min and 1 h and explored participants' experiences of using different methods in arts for health and well-being evaluation contexts, identifying strengths, challenges, limitations and gaps in evaluation practice.

We conducted two focus groups. The first (FG1) was attended by 13 participants drawn from a range of backgrounds including regional arts for health organizations, community arts groups, NHS commissioning groups, arts therapy and freelance arts evaluation practice. This group discussed the state of arts for health and well-being evaluation, practice, responding to the prompt, "What are the main challenges to evaluation of arts, health and wellbeing?" Guided by the iterative SRG learning process, the second focus group (FG2) explored the challenge of assessing value, including using methodologies of economic evaluation, in arts for health and well-being. FG2 was attended by 12 participants drawn from arts for health and well-being practice, evaluation, research, commissioning and policy-making.

The interviews and focus groups were audio recorded. The interviews were transcribed in full and subjected to thematic analysis (Braun & Clarke, 2006), undertaken by two researchers using NVivo 10. Written informed consent was obtained for all data capture activities. Project data are anonymized in all reporting. The research was approved by the UWE, Bristol Research Ethics Committee.

Survey findings

The descriptive survey was completed by 25 respondents, the majority of whom were from the original sample as few people responded to the open invitation on the website. Respondents were drawn from a range of backgrounds that reflect the target population of people with an interest in arts for health evaluation. They included arts project managers (12), researchers (10), independent evaluators (7), arts practitioners (6), commissioners and funders (7) and health care practitioners (2). While this is a voluntary and self-selecting sample, the spread of backgrounds and experiences suggests that they should be able to offer useful and relevant insights into current evaluation challenges and needs.

The survey data provide a snapshot of the range of evaluation approaches currently used in arts and health. The data are essentially descriptive and this is reflected in their analysis and presentation. Between them, respondents described 52 evaluation projects undertaken in the last five years, some of which are ongoing. The projects being evaluated spanned a range of art forms, including visual arts, music, performing and literary arts and were based in a variety of settings including museums and heritage sites, health care, communities,
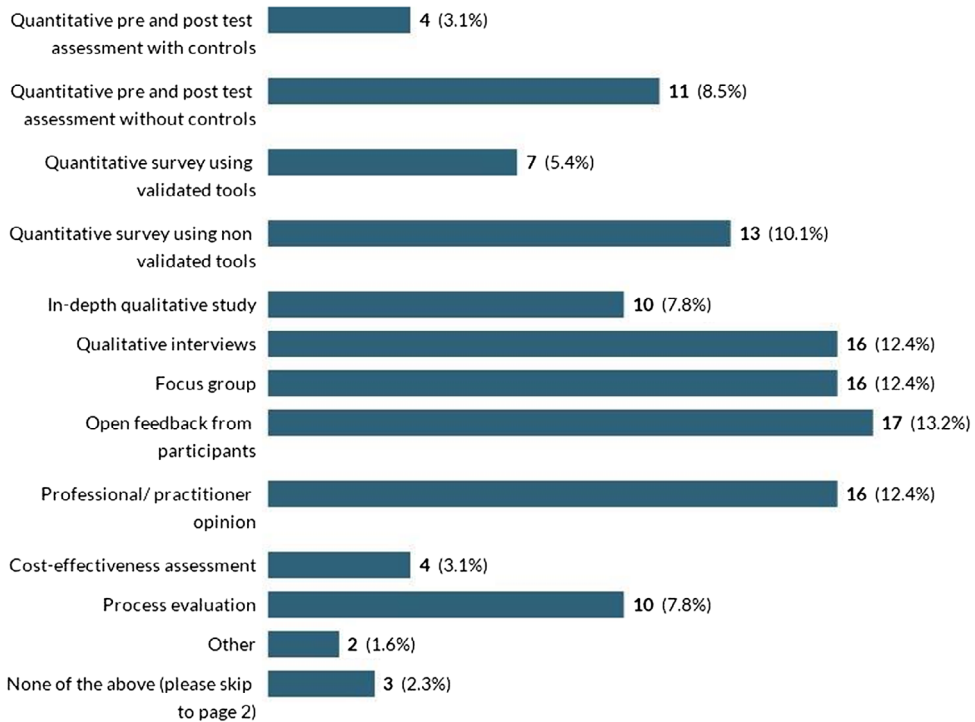
Figure 1. Arts and health evaluation approaches used in the last five years?

schools, criminal justice and military settings. Their target populations included people of all ages and people with specific conditions including mental health conditions, neurological disorders, stroke, dementia and learning disabilities. The primary funders of these evaluations were: charities (14); local authorities (12); NHS (10); UK research funding councils (4); arts organizations (4); social care providers (2); NIHR (1); and Universities (1). Some projects were brief and a few were longitudinal, with the average evaluation timescale being just under two years.

Respondents were asked whether they had, within the last five years, used any of a presented list of evaluation methodologies (Figure 1). Respondents reported the use of a wide range of evaluation methods, with extensive use of informal, "anecdotal" methods such as comments slips, feedback from artists and participants, practitioner diaries and ad hoc case studies. More formal qualitative methods such as interviews and focus groups were also reported, with quantitative methods and the use of validated assessment tools such as the Warwick-Edinburgh Mental Well-being Scale (Tennant et al., 2007) reported less frequently. Open text comments further revealed that at least four controlled studies were undertaken along with four cost-effectiveness assessments. A small number of respondents reported using physiological measures such as saliva samples as well as clinical measures of pain, cognition functioning, independent functioning and mobility.

Respondents were asked whether they had used creative- or arts-based methodologies for the purposes of evaluation, and 16 out of 25 respondents reported no usage of these methods. Examples of creative methods include offering people who struggle with language the chance to sculpt clay faces to show their responses before and after an arts project; using

music to enhance written descriptions of project outcomes; using photography to explore and document participants' experiences; and using social media to engage young people in creative evaluation.

In free text comments, respondents commented on the strengths that creative arts-based methods bring to evaluation. Arts are described as supporting reflection and deepening understanding of participants' experiences of projects:

> It provided a different language; brought to the surface ideas and perspectives that wouldn't have come through language based methods.

> It has enabled practitioners to have a window on how the intervention is or isn't impacting on participants.

> As the main point of our projects is to support participants to create original music that reflects or addresses personal challenges, recordings of the outputs often demonstrate the well-being outcomes more clearly than participants are willing to say in qualitative interviews.

They are also reported as empowering for participants, particularly those who might struggle with conventional evaluation approaches.

> Photography as an evaluative tool enriches the written evaluation and makes evaluation accessible to all. Visual art and creative writing can offer a great way for participants to evaluate their experiences. Particularly with young people, social media is the perfect evaluative tool, using platforms such as Twitter and Facebook.

Finally, creative arts are seen as assisting dissemination, making the results of evaluation more accessible to diverse audiences.

> It makes the results easier for audiences to understand and allows us the opportunity of promoting the work more widely.

Despite these advantages, the low usage of creative methods may be a consequence of perceptions of their lack of reliability and validity compared with other evaluation approaches.

> The NHS is very locked into the "evidence based practice" mantra, which traditionally favours standardized measures and RCT type methodologies. Qualitative evidence is often seen more in terms of "patient engagement" and "feedback" rather than robust evidence. It's hard to challenge this constructively.

> From a commissioner perspective, it is difficult to make informed decisions about whether to commit public monies to activity and organisations that cannot be compared as easily (or at all) on a like for like basis in the same way that a standardized questionnaire might enable comparison.

> There is something that almost borders on a kind of stigma – the term "fluffy" is often used to describe outcomes that are open to interpretation and not easily quantifiable. This can undermine the credibility and perceived professionalism of the entire project delivery and organisation, and not just the evaluation.

The questionnaire examined evaluators' experiences of working with project commissioners and funders. While positive experiences were reported, there is also some dissatisfaction with the commissioning process, with some respondents disagreeing with the statement: "I am satisfied with the process of matching needs and expectations between project evaluation and funders/commissioners" (see Figure 2). A number of challenges were identified, including managing stakeholder expectations, language and cultural differences, and methodological conventions and requirements. Commissioners and funders are often portrayed as having a powerful role in steering the evaluation process towards their preferred outcomes and methodologies, sometimes with detrimental effects. For example, a preference for

Figure 2. Satisfaction with evaluation funding/commissioning process.
Prompt: "I am satisfied with the process of matching needs and expectations between project evaluation and funders/commissioners".

quantitative methods can lead to inappropriate evaluation activities, such as ill-considered attempts at measurement attempted too early in a project, or the devaluing of impact assessment and learning about the subjective experiences of participants. The survey findings suggest that while evaluation practice is diverse, with arts-based methods sometimes adding value and richness, generally, artistic world views hold relatively little traction in a field that is dominated by health care evaluation discourse, including underlying notions of a hierarchy of evidence. Respondents stressed the importance of mutual respect by stakeholders, together with a willingness to adopt collaborative approaches, in mitigating challenges of commissioning. They also emphasized the need to work closely together at an early stage of the process in processes of co-production.

Qualitative findings

Five key themes were identified from the interviews and focus groups:

(1) Opportunities and challenges for arts and health.
(2) Evaluation frameworks, methods and tools.
(3) Issues in evaluation practice.
(4) Standardization and "scaling up".
(5) Co-production and its challenges.

Opportunities and challenges for arts and health

Participants highlighted current opportunities for arts in health and social care:

> … dementia is a bit of an open door generally because it's seen that medical solutions are not gonna work. So people in medical professions are a little bit more like, okay what are our other options. They're more … there's more receptiveness … So I think that within dementia and the arts, we may not be up against quite so many barriers. (Service provider)

> It's about gaps in the service as well, if we didn't continue to commission where would these people go? They would end up going back to primary care or being admitted or you know … I think that's why we continued to commission it. (Commissioner)

They also identified challenges, including funding for the type of evaluation that is presumed to be needed to support innovative service development.

> … Often organisations just don't have the budgets … they've got, I don't know a quarter million pounds or something but they … can only allocate about ten thousand for an external

evaluation. And it's got to include VAT. [Laughs] So you know for a five year project … (Evaluation researcher)

Evaluation frameworks, methods and tools

Commissioners need clear project descriptions that allow arts to be compared with other interventions:

> In primary care we've got one code for social prescribing, but … (that) can cover lots of stuff … so it's … how we can disentangle some of that. (Commissioner)

Commissioners also juggle competing priorities. They are required to show evidence that programmes can meet identified needs, but this isn't enough to justify financial support, they must also be confident that programmes can deliver cost-effectiveness:

> … we do look at outcomes but we have to look at the cost savings as well, because ultimately when you're commissioning services, there's only one pot of money, so if you just commission on outcomes all the time, then obviously you'll just be commissioning everything … (Commissioner)

Balancing different evaluation priorities can introduce challenges and increase demands on project evaluation. In order to be commissioned, arts organizations and practitioners may be required to show how their work "fits" within complex established systems:

> … (commissioners) have key performance indicators, they have key measures and monitoring data they collect from their providers who are delivering these services and …. we have to take that into account. (Commissioner)

There is a sense that these conventional evaluation approaches are not conducive to arts projects. For example, assessment tools that focus on negative experiences such as depression, may be detrimental in arts contexts:

> … tools exist, … you're not gonna want to sit all those people down after their music workshop with an EQ5D questionnaire and get them to fill it in before and after, how intrusive … (Service provider)

In this context, artists and arts organizations can find the demands of outcomes evaluation daunting:

> … Say you are able to focus on some specific outcomes … social inclusion, let's say, … you … agree … to measure that, and then scratch your head "How the heck am I gonna measure that?". So that's always a big issue, and whether it comes down to not having decent quantitative measures of those things, or not having the abilities or the systems to measure, analyse these things … (Service provider)

They also sometimes feel overwhelmed by unfamiliar health services culture and language:

> … I often go to events and people talk about arts for health and wellbeing and I look at them and I think, they're talking about something completely different to what I do, why are they using the same … words? (Service provider)

It was also suggested that greater rigour may be demanded from evaluation of arts programmes, which are unfamiliar to many health care providers, than is required of other, less innovative services:

> … we kind of demand a greater evidence base from them, (than) from some of our traditional providers who we probably throw millions at (Commissioner).

For some, there was a need for a more realistic approach:

> … I'd like evaluation to be more realistic … make sure that everything was evaluated appropriately according to its importance and level of resource … (Service provider)

Issues in evaluation practice

While arts are often welcomed in busy health care settings where they can add value to day to day lives of patients and carers, evaluation processes and procedures are seen as creating additional pressures and sometimes getting in the way of practice:

> … it was apparent that while arts interventions are welcomed by staff … there is less support for evaluation and research, which can be perceived as an unwelcome addition to staff workloads. (Service provider)

Evaluation is also sometimes seen as burdensome for project participants:

> … just how many times do you want to ask someone to fill in their questionnaire or answer a question or go online and tick a box? (Service provider)

Implementing evaluation can be particularly challenging for artists and facilitators who are often asked to administer questionnaires just as arts activities are getting started:

> The … facilitators weren't happy doing a load of questionnaires … the day they started … because people have come to [do arts], they haven't come to be part of a research project, which … and the facilitators were quite resistant really because I think originally … they did have a go at asking people. And they just got fed up with being told you know to bugger off. (Evaluation researcher)

A further difficulty in evaluation practice is differentiating evaluation from research. In practice, the distinction can be difficult to draw:

> … it's like the two things have got rather blurred and we're saying that the music is a service improvement because they're already providing activities, none of which have they had any medical approval for. It's like tea on a Wednesday afternoon, you're not gonna go and get ethical approval for that. Some people might not like tea (laughter), you know what I mean? (Service provider)

Some projects are steered away from definitions of research, which usually requires a process of formal ethics approval, which can be slow, cumbersome, costly and ill-suited to the requirements of arts projects:

> (The Ethics Committee) will expect to know how many we're going to investigate and what exactly … (the researcher will be) doing. So that immediately challenges that notion of very responsible, fluid methodology … the expressive way that the arts tends to relate to people, but I completely respect the basic premise around avoiding harm and making sure people know what's going on … it's tricky. (Service provider)

While commissioners are often perceived as demanding a high level of rigour, in reality they often favour more pragmatic approaches.

> … they don't want to be completely overrun with collecting data, so it's making sure it's the right data they collect. (Commissioner)

They also often work with relatively short timescales, meaning that full evaluation research is not feasible in many instances. In practice, it is not surprising that in reality many projects are evaluated using a mixed methods approach:

> … I would say 90% of our evaluations are mixed methods … even the very formative ones have a heavy qualitative element. (Commissioner)

… some will lend themselves to quantitative, some will lend themselves to qualitative depending on … the focus of the project … So you know we should celebrate plurality and inter-disciplinarity and … various methods. And experimental arts based methods, as well, you know there's not enough of that. (Commissioner)

However, with regard to qualitative methods, there is a suggestion that these are not always used with sufficient skill or rigour to produce trustworthy findings:

Yeah, so I think with qualitative stuff … I don't think there's a thematic analysis going on for qualitative feedback, but there might be some cherry picking. (Commissioner)

Standardization and "scaling up"

Across the data there is some support for the idea that sector development might be strengthened by the adoption of a standardized evaluation framework:

… I know in other areas … they have … standard evaluation frameworks for obesity, healthy eating, physical activity … it was quite useful (to have) just a checklist of what things you should collect. (Commissioner)

Using clear evaluation frameworks can bring clarity to project planning:

… but one of the things, so I did use logic modelling … to construct an Evaluation Framework … I really think that helped … us identify the clear measures, the indicators, and therefore, focus the evaluation. (Service provider)

However, flexible frameworks are needed if they are to apply across the diverse arts and health sector:

The sector is vast and I think that's a really, really big challenge because how can you come up with one particular tool for … a public art project in a hospital, a mental health based project and a smoking behaviour change project and they're all arts for health and wellbeing projects, but … the methodologies, the approach, the skills, they are all so different. (Service provider)

There is a danger that standardized frameworks may limit the rich diversity that is a strength of the sector:

We don't want uniformity because that can be quite stultifying. (Commissioner)

Yeah but equally you want to produce a tool that doesn't in any way affect their unique support they offer … We don't want to disturb their (arts) outcomes of what they can deliver in their uniqueness of their service. (Commissioner)

As well as standardization, the notion of, "scaling up" was discussed. Rather than emphasizing the uniqueness of what arts projects offer, some stakeholders advocated developing larger, replicable interventions as well as investing in networking and infrastructure development underpinned by research approaches such as economic evaluation to consolidate the arts and health sector. Others disagreed, arguing that arts projects have little experience of scaling up and that they are not generally in a strong position: failed attempts to go to scale would be disempowering and could lead to a loss of focus on core values.

Co-production and its challenges

A key theme, particularly emphasized by commissioners, is that of co-production, usually meaning consultation, collaboration and the development of a shared vision between commissioners, arts delivery organizations, participants and evaluators. Co-production was

advocated as the strategy for overcoming some of the evaluation challenges that participants encountered. Co-production seems particularly important during evaluation planning in order to fully embed evaluation in service delivery. The implication of this is that longer lead in times and more extensive consultation about methods are needed than are currently the norm.

> I do feel that that makes the evaluation stronger because then you get better buy in … if you design it on your own, you're never gonna satisfy everyone about whether they feel engaged or involved in it enough … you can do your best but I do think those that are co-produced and developed together are more … successful because you've got buy in from the right people so the right people are there. (Commissioner)

Co-production allows evaluation to draw on appropriate knowledge and expertise:

> it's … about … who do you need round the table, and sometimes it might be the evaluation experts working alongside the artists who really understand what they're delivering … it needs to be the right clinician, it needs to be a public health specialist … because obviously they're brilliant at helping identify the right outcomes, so yeah, it's about having the right people round the table and I think, I think sometimes … there is still an element of, "oh I'll just do it myself, I'll just write it (laughs)". (Commissioner)

Co-production may also have the advantage of helping to focus evaluation and can prevent evaluation from setting off or proceeding on the wrong track:

> I think part of the problem is around the university, which evaluated it, I don't think they were given a clear steer, so … is all great and nice but that doesn't satisfy all commissions. (Commissioner)

Co-production is seen as a way of ensuring that no one gets left behind and that different people's skills and expertise are valued. This can enhance evaluation practice by ensuring that the backgrounds, knowledge and skills of people who are to be relied on to deliver the evaluation are considered:

> … if you're expecting people to collect data, like the artists on the front line, well then they need to engage in the development, cos actually they might already use (a tool) so they're already skilled in applying it or you, cos you might find you're introducing … an alternative scale, and they're like "oh I've never used this before, I've no idea how to use this", and you're almost falling at the first hurdle … (Commissioner)

Finally, co-production is seen as enhancing the quality of participants' experience of evaluation, making them more likely to engage in it in the future and underpinning stakeholder satisfaction with the outcomes:

> … I think that's a key to a successful evaluation, that I don't come as an expert … if we can make it a meeting of minds and also if we can involve participants as well in that process. So I think some of my best work has been like that. There's been genuinely participatory … And then you're more likely to get a finished product that everybody's happy with. (Evaluation researcher)

Discussion

The participants in this study were keenly aware of the current landscape of opportunities for arts-based approaches to address gaps in health and social care provision. Along with opportunities, some participants identified significant structural and cultural barriers, with the arts sector portrayed as relatively weak, being forced to adopt approaches and fit into frameworks that are inflexible and insensitive to their values and contexts. Those from arts backgrounds can feel disadvantaged, coming from a sector that is relatively fragile and has

a poorly developed infrastructure compared with other sectors, and which has been de-prioritised in recent rounds of austerity-driven spending cuts.

Evaluation can play a critical role in developing innovative and robust services that strongly engage members of the public. However, a number of barriers to evaluation were identified. Problems can arise when commissioners, funders and practitioners and evaluators have different understandings of the purposes of evaluation. Across the data there is a general perception that commissioners seek robust evaluation designs that can deliver outcomes evidence, perhaps modelled on clinical research. For some, the balance has shifted too far towards outcomes evaluation, with a consequent devaluing of other forms, such as process evaluation and reflective practice that is seen as throwing arts projects off track. Hence, there is a view that artists should stay true to their artistic aims in the face of pressure to fit into frameworks imposed from outside. However, this is allied with a recognition that evaluation often struggles to capture the elusive "essence" of arts, or fails to show why participation in creative arts is different from other activities, such as walking or gardening. As well as being frustrating for service providers, this makes it difficult for commissioners to understand how arts can fit within a complex landscape of services that could potentially be supported.

The data suggest that artists may be disadvantaged by a lack of access to evaluation knowledge and resources, and that there is at times an "un-level playing field", with a suggestion that innovative services face a higher burden of evaluation than established programmes. Unequal power relationships may also account for the low usage of arts-based methodologies, despite widespread acknowledgement of the richness that they can bring. These methods, often participatory in nature, are subject to similar criticisms regarding bias and the blurring of boundaries between the researcher and researched that are levelled at participatory methods of enquiry more generally (Fraser & al Sayah, 2011).

Given the ambivalence about conventional forms of evaluation within the arts sector, it is tempting to argue that advocacy is a more effective tool for service development. Some of our respondents, like those in Goulding's study of stakeholder perspectives on evidencing impacts of art outcomes, view evaluation as a burden. The suggestion is that, in terms of influencing policy and funding bodies, well-focused advocacy may be more effective than attempts to generate more evidence (Goulding, 2014). However, most participants in our study recognized the limits of advocacy that is not supported by evidence. Rigorous evaluation serves as a reminder that arts can have not just benign effects: evaluation is needed in order to understand what doesn't work as well as what works.

There is a suggestion from the data of a need for a standard evaluation framework in arts, health and well-being, provided that this allows space for uniqueness and creativity. Greater standardization may encourage practitioners to work together and allow projects and programmes to share findings, learning and experiences. However, such a framework would need to accommodate a diverse array of methodologies and outcomes, which range from individual health and well-being outcomes to social impacts such as reducing inequalities and promoting cultural change. A framework based on the MRC complex interventions evaluation framework (Craig et al., 2008) has been carefully applied to the arts and health context (Fancourt & Joss, 2014). However, the field may not be quite ready for what is sometimes perceived as an imposition of medically based hierarchies of evidence. Since this study was completed, The PHE Arts, Health and Well-being Evaluation Framework has been published by Public Health England (Daykin with Joss, 2016). This document, developed by

Aesop and The University of Winchester, UK, offers a reporting tool modelled on standard public health frameworks (Cavill, Roberts, & Rutter, 2012; Roberts, Cavill, & Rutter, 2012). The tool encourages transparency whilst allowing for pragmatic and diverse evaluation approaches. Its adoption may help to strengthen the visibility of the sector and would support practitioners and commissioners in understanding the similarities and differences between projects as well as how they can contribute to health and well-being.

Guidance may also be needed on ethical evaluation. Ethical requirements can increase the evaluation burden, and the requirements of formal ethical approval and research governance can overwhelm small projects. Consequently, projects are often steered away from the label, "research". While this offers short term advantages, it means that evaluation protocols do not receive the benefit of peer review and are less likely to produce findings and learning that can be disseminated beyond the immediate practice context. Perhaps there is an additional need for proportionate methods of providing peer assessment to ensure ethical practice in evaluation.

Developing effective co-production was favoured, especially by commissioners, as the strategy for overcoming many evaluation challenges. The NESTA definition of co-production is often adopted in discussions about evaluation:

> Co-production means delivering public services in an equal and reciprocal relationship between professionals, people using services, their families and their neighbours. (Boyle & Harris, 2009)

While co-production emerged as a key theme across the data, many challenges and barriers to co-production are revealed in this study. Some of these originate within the arts sector, where involvement of both a commissioning and evaluation partner within the co-production relationship is relatively new. This sector is to some extent limited by structural constraints, being made up of small, independent organizations that are often competing with each other for scarce resources. Other barriers emanate from differences in status, power, language, values and culture between stakeholders, with commissioners and funders seen as sometimes controlling and shifting the goal posts regarding requirements of rigour and evidence standards. These barriers can get in the way of effective co-production, diminishing the voices of those who are less powerful in the process, including service providers and service users.

Conclusion

Arts organizations face increasing challenges to show outcomes and cost-effectiveness of their work and there is a need to develop appropriate evaluation models and frameworks as well as capacity and resources for the sector. This knowledge exchange project sought to examine experiences of evaluation of arts and health interventions from different standpoints. The findings draw on the views of a focused sample of stakeholders including artists, health professionals, evaluators, researchers, project managers, commissioners, funders and policy-makers. They encompass wide ranging experience of evaluation of diverse arts practice, from visual and performing arts to culture and heritage, used in public, private and third sector health and well-being organizations.

The study suggests that the changing landscape of health and social care offers genuine opportunities for the development of innovative, arts-based service models that can address needs and improve services. However, there is some ambivalence from the arts sector about how to engage with these opportunities. Commissioners require data that will demonstrate

both effectiveness and cost-effectiveness in order to justify funding services. This triggers concerns about the appropriateness of evaluation methodologies and hierarchies of evidence, underlined by an awareness of power relationships, with participants from the arts sector often feeling disadvantaged in discussions.

Amongst some experienced evaluators there is a view that the commissioning process can be made difficult by conflicting and unclear expectations as well as differences in language, culture and methodological conventions. There is a suggestion that the process steers evaluation towards particular frameworks and methodologies and that the domination of health care evaluation discourse, including underlying notions of a hierarchy of evidence, can be detrimental to projects and participants. The lack of a shared vision amongst stakeholders about the purposes of evaluation, together with low priority given to evaluation resourcing, the weight of research governance and the lack of appropriate frameworks can make evaluation confusing and daunting for practitioners. It is difficult for arts organizations in this context to develop robust responses to requirements to demonstrate project outcomes.

While commissioners and funders are sometimes portrayed as steering the evaluation process towards their preferred outcomes and methodologies, sometimes with detrimental effects. The commissioners who took part in this study inhabit a more pragmatic world view. This suggests that commissioner-led evaluation can allow for a range of methods and can accommodate proportionate assessment and diverse methods. Nevertheless, there is a need for demonstrable rigour in evaluation, especially in regard to findings from qualitative and arts-based methodologies, which are more readily dismissed than other types of data as being too closely aligned with advocacy claims.

Participants in this project acknowledge that co-production between commissioners, arts delivery organizations, participants and evaluators, is needed at each stage of the evaluation cycle. Effective co-production takes time and resources in order to ensure that no-one is left behind or devalued by the evaluation process. Co-production of evaluation is not well established within the arts sector, where small organizations often face fierce competition for scarce resources. Nevertheless, this notion seems to offer a genuine way forward in aligning arts perspectives with those of health and well-being, with the potential to support innovative, high-quality, cost-effective services.

Co-production may require changes in evaluation practice, such as extending planning timescales and using effective consultation methods, which in turn require realistic resourcing. Even with these changes, there may be barriers to co-production emanating from the inherently fragile nature of the arts sector as well as from differences in status, power, language, values and culture between stakeholders. These might be addressed to some extent by training and professional development for those who are relatively new to evaluation. The Creative and Credible Website (www.creativeandcredible.co.uk) that was produced as part of this project seeks to support such initiatives by providing freely available resources. Beyond this, strategies and resources are needed to ensure that those seeking to develop arts for health and well-being are able to access evaluation knowledge and expertise, develop appropriate frameworks and tools, engage with commissioning, funding and policy agendas, reflect on participants' experiences of arts projects and learn from the process while remaining true to their artistic purposes.

Strengths and limitations of the study

The study draws on a focused sample to include a range of perspectives and experience. While it cannot claim to include every instance of arts for health evaluation, the project draws on the views of a wide range of stakeholders who reflect a range of roles and perspectives across the field of arts and health evaluation practice. As a knowledge exchange project of one year's duration we were not able to fully engage in extended research. The project is limited to a UK sample, however, the evaluation challenges we discuss are likely to affect many contexts where resources are scarce and there is pressure to demonstrate outcomes and cost-effectiveness when justifying expenditure on the arts. The Creative and Credible Website has already received positive feedback from users not confined to the UK. A useful venture for further research would be to examine similarities and differences across international evaluation contexts.

Acknowledgements

Funding

ORCiD

Norma Daykin http://orcid.org/0000-0002-1114-6780

References

Baird, A., & Samson, S. (2015). Music and dementia. *Progress in Brain Research, 217*, 207–235. doi:10.1016/bs.pbr.2014.11.028

Boyle, D., & Harris, M. (2009). *The challenge of co-production*. London: Nesta.

Braun, V., & Clarke, C. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77–101. doi:10.1186/1477-7525-5-63

Bungay, H., & Clift, S. (2010). Arts on prescription: A review of practice in the UK. *Perspectives in Public Health, 130*, 277–281. doi:10.1177/1757913910384050

Camic, P., Tischler, V., & Pearman, C. H. (2014). Viewing and making art together: A multi-session art-gallery-based intervention for people with dementia and their carers. *Aging & Mental Health, 18*, 161–168. doi:10.1080/13607863.2013.818101

Camic, P. M., Williams, C. M., & Meeten, F. (2013). Does a "singing together group" improve the quality of life of people with a dementia and their carers? *A Pilot Evaluation Study. Dementia, 12*, 157–176. doi:10.1177/1471301211422761

Carnwath, J. D., & Brown, A. S. (2014). *Understanding the value and impacts of cultural experiences: A literature review*. Manchester: Arts Council England. Retrieved April 23, 2016, from http://www.artscouncil.org.uk/publication/understanding-value-and-impacts-cultural-experiences

Cavill, N., Roberts, K., & Rutter, H. (2012). *Standard evaluation framework for physical activity interventions*. Oxford: National Obesity Observatory. Retrieved April 23, 2016, from www.noo.org.uk/core/frameworks/SEF_PA

Chatterjee, H. J., & Camic, P. M. (2015). The health and well-being potential of museums and art galleries. *Arts & Health, 7*, 183–186. doi:10.1080/17533015.2015.1065594

Clift, S. (2012). Creative arts as a public health resource: Moving from practice-based research to evidence-based practice. *Perspectives in Public Health, 132*, 120–127. doi:10.1177/1757913912442269

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008, September). Developing and evaluating complex interventions: The new medical research council guidance. *BMJ, 337*, a1655. doi:10.1136/bmj.a1655

Cusack, P., Lankston, L., & Isles, C. (2010). Impact of visual art in patient waiting rooms: Survey of patients attending a transplant clinic in Dumfries. JRSM Short Rep, *1*, 52, 1–5. doi:10.1258/shorts.2010.010077

Daykin, N., Attwood, M., & Willis, J. (2013). Supporting arts and health evaluation: Report of a UK knowledge transfer partnership. *Journal of Applied Arts in Health, 4*, 179–190. doi:10.1386/jaah.4.2.179_1

Daykin, N., Byrne, E., Soteriou, T., & O'Connor, S. (2010). Using arts to enhance mental healthcare environments: Findings from qualitative research. *Arts and Health, 2*, 33–46. doi:10.1080/17533010903031408

Daykin, N., de Viggiani, N., Pilkington, P., & Moriarty, Y. (2013). Music making for health, well-being and behaviour change in youth justice settings: A systematic review. *Health Promotion International, 28*, 197–210. doi:10.1093/heapro/das005

Daykin, N., Orme, J., Evans, D., Salmon, D., McEachran, M., & Brain, S. (2008). The impact of participation in performing arts on adolescent health and behaviour: A systematic review of the literature. *Journal of Health Psychology, 13*, 251–264. doi:10.1177/1359105307086699

Daykin, N., Willis, J., Gray, K., & McCree, M. (2016). *Creative and credible: Developing evaluation resources for arts, health and wellbeing. Report of a knowledge exchange project*. Bristol: UWE. ISBN 9781860435249 (1860435246).

Daykin, N. with Joss, T. (2016). *Arts, health and wellbeing, an evaluation framework*. Public Health England Publications. Gateway number 2015595. Retrieved from https://www.gov.uk/government/publications/arts-for-health-and-wellbeing-an-evaluation-framework

Fancourt, D., & Joss, T. (2014, July). Aesop: A framework for developing and researching arts in health programmes. *Arts & Health, 7*, 1–13. doi:10.1080/17533015.2014.924974

Fraser, A., Bungay, H., & Munn-Giddings, C. (2014). The value of the use of participatory arts activities in residential care settings to enhance the well-being and quality of life of older people: A rapid review of the literature. *Arts & Health: International Journal for Research, Policy & Practice, 6*, 266–278. doi:10.1080/17533015.2014.923008

Fraser, K. D., & al Sayah, F. (2011). Arts-based methods in health research: A systematic review of the literature. *Arts & Health: International Journal for Research, Policy & Practice, 3*, 110–145. doi:10.1080/17533015.2011.561357

Gelo, F., Klassen, A. C., & Gracely, E. (2015). Patient use of images of artworks to promote conversation and enhance coping with hospitalization. *Arts & Health, 7*, 42–53. doi:10.1080/17533015.2014.961492

Goulding, A. (2014). Arts on prescription for older people: Different stakeholder perspectives on the challenges of providing evidence of impact on health outcomes. *Journal of Applied Arts & Health, 5*, 83–107. doi:10.1386/jaah.5.1.83_1

Gregory, H. (2011). Using poetry to improve the quality of life and care for people with dementia: A qualitative analysis of the try to remember programme. *Arts and Health, 3*, 160–172. doi:10.1177/1471301215611763

Johnson, J., Culverwell, A., Hulbert, S., Robertson, M., & Camic, P. M. (2015). Museum activities in dementia care: Using visual analog scales to measure subjective wellbeing. *Dementia*. First Published Online, October 13 2015. doi:10.1177/1471301215611763

Kreutz, G., & von Ossietzky, C. (2015). The value of music for public health. In S. Clift & P. Camic (Eds.), *Oxford textbook of creative arts, health, and wellbeing: International perspectives on practice, policy and Research* (pp. 211–219). Oxford: Oxford University Press.

McDaid, D., & Park, A. (2013). *Investing in arts on prescription: An economic perspective (project report)*. London School of Economics [online]. Retrieved October 23, 2014, from http://artsandminds.org.uk/reports-and-documentation/

Mowlah, A., Niblett, V., Blackburn, J., & Harris, M. (2014). *The value of arts and culture to people and society*. Manchester: Arts Council England. Retrieved February 27, 2015, from www.artscouncil.org.uk/media/uploads/pdf/The-value-of-arts-and-culture-to-people-and-society-An-evidence-review-TWO.pdf

Nanda, U., Chanaud, C., Nelson, M., Zhu, X., Bajema, R., & Jansen, B. H. (2012). Impact of visual art on patient behavior in the emergency department waiting room. *The Journal of Emergency Medicine, 43*, 172–181. doi:10.1016/j.jemermed.2011.06.138

Osman, S. E., Tischler, V., & Schneider, J. (2014). "Singing for the brain": A qualitative study exploring the health and well-being benefits of singing for people with dementia and their carers. *Dementia*. First published Online November 24, 2014. doi:10.1177/1471301214556291

Renton, A., Phillips, G., Daykin, N., Yu, G., Taylor, K., & Petticrew, M. (2012). Think of your art-eries: Arts participation, behavioural cardiovascular risk factors and mental well-being in deprived communities in London. *Public Health, 126*, S57–S64. doi:10.1016/j.puhe

Rickett, C., Greive, C., & Gordon, J. (2011). *Something to hang my life on*: The health benefits of writing poetry for people with serious illnesses. *Australasian Psychiatry, 19*, 265–268. doi:10.3109/10398562.2011.562298

Roberts, K., Cavill, N., & Rutter, H. (2012). *Standard evaluation framework for dietary interventions*. Oxford: National Obesity Observatory. Retrieved February 27, 2016, from www.noo.org.uk/core/frameworks/SEF_Diet

Simpson, K., & Keen, D. (2011). Music interventions for children with autism: Narrative review of the literature. *Journal of Autism and Developmental Disorders, 41*, 1507–1514. doi:10.1007/s10803-010-1172-y

Skingley, A., Bungay, H., & Clift, S. (2011). Researching participatory arts, wellbeing and health: Methodological issues. *Journal of Arts and Communities, 3*, 73–87. doi:10.1386/jaac.3.1.73_1

Skingley, A., Page, S., Clift, S., Morrison, I., Coulton, S., Treadwell, P., … Shipton, M. (2014). Singing for breathing: Participants' perceptions of a group singing programme for people with COPD. *Arts & Health, 6*, 59–74. doi:10.1080/17533015.2013.840853

Sonke, J., & Lee, J. B. (2015). Arts for health in community settings: Promising practices for using the arts to enhance wellness, access to healthcare, and health literacy. In S. Clift & P. Camic (Eds.), *Oxford textbook of creative arts, health, and wellbeing: International perspectives on practice, policy and research* (pp. 103–212). Oxford: Oxford University Press.

Sonke, J., Pesata, V., Arce, L., Carytsas, F. P., Zemina, K., & Jokisch, C. (2015). The effects of arts-in-medicine programming on the medical-surgical work environment. *Arts & Health, 7*, 27–41. doi:10.1080/17533015.2014.966313

Staricoff, R. (2004). *Arts in health: A review of the medical literature*. London: Arts Council England. Retrieved February 27, 2016, from www.artscouncil.org.uk/publication_archive/arts-in-health-a-review-of-the-medical-literature

Staricoff, R., & Clift, S. (2011). *Arts and music in healthcare: An overview of the medical literature: 2004–2011*. London: Chelsea and Westminster Health Charity [online]. Retrieved February 27, 2016, from http://www.publicartonline.org.uk/whatsnew/news/article.php/Arts+and+Music+in+Healthcare%3A+An+overview+of+the+medical+literature%3A+2004-2011

Tennant, R., Hiller, L., Fishwick, R., Platt, P., Joseph, S., Weich, S., … Stewart-Brown, S. (2007, November 27). The Warwick-Edinburgh mental well-being scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes*. doi:10.1186/1477-7525-5-63

Zhang, J. M., Pu, W., Yao, J., Zhao, L., Davis, M. P., Walsh, D., & Yue, G. H. (2012). Music interventions for psychological and physical outcomes in cancer: A systematic review and meta-analysis. *SpringerLink, 20*, 3043–3053. doi:10.1007/s00520-012-1606-5