Cultural Value

Reviewing art therapy research: A constructive critique

Shona Kelly, Larissa Davies, Deborah Harrop, Alex McClimens, David Peplow, Nicholas Pollard
Reviewing art therapy research: A constructive critique

Shona Kelly; Larissa Davies; Deborah Harrop; Alex Mcclimens; David Peplow; Nicholas Pollard

Executive Summary

The literature search that informed our review initially yielded 12,122 papers of potential interest, derived from seven databases. After applying a series of filters we arrived at 92 papers on which we base our findings, thoughts and recommendations for future work.

Our methodological approach was informed by the systematic review guidance published by the Centre for Reviews and Dissemination (2009), and the Arts Council definition of 'arts activities'. Hence we considered papers reporting therapeutic arts interventions conducted on 'patients' which included some measurement of a health state. After excluding any research on people less than age 18, we selected studies where participants had active (as opposed to passive) engagement with the therapy/treatment/medium. Only study types which were quantitative were included in this review.

Rather than simply criticise the execution of the research we applied our own expertise to the process. It was immediately evident that definitions and categories would pose some difficulties as there is much variety in the language used to describe the arts, therapies and treatment. This is a problem of indexing, causing the literature search and initial screening to be a laborious process.

The most commonly reported art activities were: writing, music, art and dance. The most numerous health condition studied was mental health followed by cognitive function, stress and cancer. Most research was carried out in the US and the UK. As a discipline, psychology featured regularly. When arts therapists were involved in the research the descriptions and possible effects of the art medium tended to be better elucidated.

Future research into the use of art therapy in healthcare will benefit from a synthesis of approaches that can retain the more robust aspects of, for example, RCTs with the insights that can be derived from qualitative methods.

Key words:

art activity, art therapy, review, evidence-based treatment, healthcare, public health, therapeutic
Reviewing art therapy research: A constructive critique

There is a long history of academic and evaluation research into health and the arts and culture, both within the UK and in other countries (Fraser and al Sayah, 2011). This evidence includes individual impacts, covering therapeutic and clinical outcomes for patients, and broader community impacts (e.g. reduced health care needs) in both clinical and non-clinical populations.

Within the body of research on arts and healthcare, there have been attempts at measuring and valuing the effects of the arts on clinical outcomes, although often this is context specific (e.g. for specific amenities or initiatives/programmes in specific locations) and lacking a policy purpose. This project was an extensive review of the literature on the use of arts in therapy. In this paper the use of 'AT' refers to a range of arts applications in therapy. This report focuses on the quantitative components of this literature.

Introduction

In recent years there have been a number of reviews of art therapy. An example is Lelchuk Staricofs’ (2004) systematic review of the medical literature on arts and health. This was one of the first robust studies to carry weight with those responsible for delivering health care and several other reviews have followed (Fraser and al Sayah, 2011; Clift, 2012; Stuckey and Nodel, 2010). The reviews vary in their remit and may focus on therapeutic art; however these reviews may also examine research about the people who use art in treatment (Fraser and al Sayah, 2011) or focus on building a case for the value of art in health rather than focus on critically reviewing the evidence (Clift, 2012; Stuckey and Nodel, 2010). Across the reviews, little of the research examined was quantitative, which is a potential problem as quantitative methods can provide a more robust evidence of benefit.

The one consistent feature across the reviews was a statement about the clear need for more research to map out the terrain. For example, when considering the use of art in health research, Fraser and al Sayah (2011) observe that despite a long history of use, 'There is no clear understanding of the kind of arts or the way the arts have been used as a research method' (Fraser and al Sayah, 2011: 110). From their review they concluded that three key areas needed further study: ‘trends in arts-based methods in health research, methodological issues and theoretical issues' (Fraser and al Sayah, 2011:138). These findings are most recently evidenced by a systematic review commissioned by the UK Department of Culture Media and Sport (DCMS, 2014) as part of the CASE (Culture and Sport Evidence) programme on the social benefits of culture (and sport). That study, undertaken by one of the co-investigators of the present review (L. Davies), found evidence that the arts and culture are beneficial to both mental and physical health. Furthermore, the study found that arts and culture can be used directly to improve clinical outcomes, but also indirectly as a powerful force in the re-integration into society, which creates therapeutic benefits though social interactions with others and development of skills,
learning and other competencies. Both of these outcomes following engagement with arts and culture can bolster confidence, self-esteem and self-efficacy. However, the DCMS study, like all the other reviews, found a lack of robust evidence to support these findings, with much research based on cross-sectional studies, qualitative research and narrative reviews, all of which are considered weaker forms of research evidence.

Evidence-based Treatment
The lack of quality research in arts therapy studies is not an unusual situation, and is similar to the situation that most medical research was in a decade or two ago. There are now evidence hierarchy systems which classify research by study design and Figure 1 shows the evidence hierarchy most commonly applied to medical research.

Figure 1: Evidence Hierarchy System

![Evidence Hierarchy System](http://www.mstrust.org.uk/professionals/information/wayahead/articles/16032012_08.jsp)

At the top lies systematic reviews which aggregate a number of studies that all use the same research design, usually randomised controlled trials or case-control studies. Below this are randomised controlled designs which are deemed to constitute a high level of evidence;
however, RCTs are usually conducted on highly motivated people who are carefully scrutinised and so may behave in a different way under study conditions. Below RCTs are the observational study designs which were largely developed to address ethical and practical issues in healthcare research. For example, it is unethical to give people a substance known to be toxic or hazardous so observational designs exist to study people who are exposed to these substances unintentionally in everyday life. Besides study design, research needs to be evaluated on the quality of the study execution. For each type of study design there are quality criteria that include information such as number of participants, the power of the study to detect a difference, sources of bias (non-random error), and so on. These are not elitist classification systems, but are rather grounded in a genuine attempt to identify whether the therapy is having an effect or whether an improvement is just the usual progression of the disease.

There is a perception that AT does not lend itself to quantitative research techniques. This is not necessarily true because the use of quantitative methods refers to the measurement of effects, not to measurement of the treatment or therapy being used. It should not be assumed that it is impossible to employ quantitative measures when it is possible to generate appropriate measurement systems. Measures can be developed for broader social units such as reducing social isolation or improving communication.

Art therapy is not alone in having a dearth of high quality evidence and alternative approaches to a standard practice of ‘systematic review’ have been developed. This is particularly true for ‘complex’ public health interventions. The reasons people smoke, for example, are multi-factoral and it becomes difficult to determine whether a single intervention to help people quit is actually effective. Within public health the entire body of evidence is often used to support an approach in ‘realist reviews’ (Pawson et al. 2005). These broad critical reviews examine the entire body of research.

A problem arises from the complexity of providing a therapeutic activity and producing the research to evaluate a therapy. The two systems need to be run in parallel. We recognise that in some services or situations the complexity of the service conspires against working beyond the therapeutic process. But with new economic imperatives robust evaluation is too important to be a post-hoc process. There are some standard processes which we should be striving to achieve that would strengthen any evaluation and create a stronger body of evidence.

**The Therapeutic Mechanisms of Art Therapy**

An additional issue with therapeutic art is the underlying therapeutic mechanism. Multiple mechanisms have been proposed. For example: the process of sublimation in the use of arts to represent concerns; engaging with particular techniques or media; the use of the body; acquisition of skills; participation in performance or events; as well as particular approaches to the underpinning therapeutic theory which may themselves derive from many sources across psychology to folklore, cultural and anthropological knowledge with both individuals and groups across all manner of health conditions including preventative health. Is the benefit gained from the interaction with the therapy provider, interaction with others, actual physical or neurological training or the learning of new practices and recognition of abilities? Does benefit derive from the use of specific approaches, personal goals which it allows people to
accomplish or simply the diversions it creates for the people involved? In health there is a more recent recognition that non-clinical, psychosocial factors are key factors in health. The models that we use to ‘explain’ how people become ill and how healthcare provision improves health have been evolving. New models that incorporate social interaction with the therapy provider or other patients need to be examined more robustly and art therapy is particularly well placed to address this issue.

**Aims and Objectives of This Project**

A. We set out to provide the most in-depth, multi-dimensional, critical review to date of the published art therapy literature. We focused on publications that:

- were written for a healthcare audience (practitioner or healthcare planner as this audience was selected as it is the most likely to commission art therapy for treatment);
- had the patient engage in an active way with the art. For example, playing music or singing rather than passively listening to music. In part this decision was made because passive music therapy had been the subject of several previous systematic reviews (e.g. Bradt 2013);
- had a stated goal of providing evidence of therapeutic benefit from an artistic activity through measurement of a health state;
- were quantitative in nature as the evidence hierarchies used by healthcare planners place considerable emphasis on quantitative research.

B. We also set out to provide a constructive critique in plain language to describe how evaluations of art therapy treatments could be improved.

**Methods**

This review synthesises the literature on health and therapeutic art interventions. The methodological approach was informed by the systematic review guidance published by the Centre for Reviews and Dissemination (2009) and Moher et al. (2009).

Starting with the Arts Council definition of arts activities we included research that assessed **active** engagement with one of the following activities: dance for artistic and social purposes (not for fitness); singing including opera; playing a musical instrument; writing any music; performing in a play or drama; painting; drawing; printmaking; sculpture; photography as an artistic activity; making films or videos as an artistic activity; creating original artworks or animation using a computer; fibre arts, such as embroidery; woodworking, such as carving; calligraphy; pottery; jewellery making; writing stories or plays; writing poetry.

**Search strategy**

The search comprised three facets with terms relating to: (1) health, (2) therapeutic arts and, (3) to exclude specific frequently occurring aspects; for example research carried out on mice. Terms included in facet one and facet two were searched for in the title and abstract fields, whereas facet three looked for terms which occurred in all fields or a specific journal title.
Controlled vocabulary terms were used where available. No date limits were applied and only English language publications were sought.

The following seven databases were searched in this literature review: ASSIA (ProQuest, The Cochrane Library (Wiley), MEDLINE (EBSCO), PsycINFO (ProQuest), Scopus (Elsevier), Sociological Abstracts (ProQuest) and Web of Science (Thomson Reuters). These databases were selected as their scope matches the subject of this review, or in the case of Scopus and Web of Science, they are multi-disciplinary. The literature searches were undertaken in March 2014.

The literature search, as undertaken in MEDLINE, is located in Appendix 1, alongside a table showing the search numbers yielded from each database. RefWorks, a bibliographic management tool, was used to organise the literature yielded for this review. All papers not held at Sheffield Hallam University were requested from The British Library.

A limitation of this literature review is that all relevant sources of information may not have been identified, in particular grey literature. In addition, author, citation and reference searches were not undertaken.

**Study selection**

In total, the searches identified 17,061 papers which, after duplicate removal, resulted in 12,122 unique papers of potential interest. Using pre-defined criteria, all papers were independently assessed for eligibility for inclusion in the literature review. After double screening at title and abstract level, 324 papers of possible interest remained. The full texts of the remaining papers were single screened for inclusion in the review. After examination of the full texts, 232 further papers were excluded, leaving 92 papers for inclusion in this literature review.

This review included published journal papers reporting therapeutic art interventions which included a measurement of a health state. Only study types which were quantitative were included in this review. The literature search screening process is summarised in Figure 2.

**Comments on the literature search process**

The literature searches, having yielded in excess of 12,000 papers, suggest that a wide variety of terms are used to describe therapeutic art interventions and that these terms are used inconsistently. The consequence of this terminological inconsistency was that we developed a search strategy with high sensitivity which sought to capture all relevant papers and also accept that there would be low specificity (or precision). To illustrate this, the percentage of papers retained after title and abstract screening represents 2.67% of the total number of papers yielded (after duplicate removal), and the percentage of papers included in this review was less than 1% (see Figure 2, below).

As study type was attributed at point of screening (either at title and abstract or full text level), of the 11798 studies which were not included in this review, 728 were classified as either review papers, qualitative studies, case studies, background information, books or book chapters or papers containing potentially relevant references. These papers will be drawn on for future research but did not meet the remit of this project. From a medical research
The perception of a poor body of literature may result from a limited number of standard indexing terms than is available in other, more clinically focused, healthcare activities. The US National Library of Medicine provides the standardized (MeSH) terms.

**Figure 2. Flow chart of literature review selection process** (adapted from Moher 2009)

- Papers identified through database searching ($n = 17061$)
- Papers after duplicates removed ($n = 12122$)
- Papers double screened by title and abstract ($n = 12122$)
- Full-text papers single screened assessed for eligibility ($n = 324$)
- Papers including in this review ($n = 92$)
- Papers excluded, most because they were not art-based ($n = 11798$)
- Papers excluded for the following reasons: passive therapy, not healthcare activity, not therapeutic arts, no health measures, book or book section, self-help, study design description, opinion piece, qualitative, case study, background piece, review article, unable to obtain copy from The British Library, PhD thesis not available electronically, not English.
appropriate ones are: art therapy, bibliotherapy, dance therapy, music therapy, sensory art therapies but these do not seem to have been applied with consistency. It is important to understand that other research literature databases have no standardised indexing system at all. Thus in order to find the literature the group met to develop a detailed list of search terms.

The Team
This is an innovative review on using art for therapeutic processes, it combines experts drawn from disciplines that do not usually work together, thus providing a more sophisticated typology of the research than has been published previously. Shona is an expert in the critical review of literature and the non-clinical determinants of health as well as being trained in interdisciplinary research. She has also used fibre arts as a personal process for managing stress and traumatic life events. Nick is a qualified occupational therapist who teaches and uses art in practice. Larissa has experience of researching the impacts and value of arts and culture. She also conducted a recent systematic review on the social benefits of culture, which included an examination of arts and health literature. David's doctoral research was funded by the AHRC. He is a specialist in linguistic approaches to social interaction and this approach is key for teasing out the, often unstated, mechanisms presumed to lie behind the art therapy research. Deborah is an Information Scientist working in the health, social care and biosciences subject areas. She is an expert at designing and undertaking complex multi-faceted literature searches. Alex is a Research Fellow in Learning Disability Nursing with a background in participatory research methods.

Research Protocol
We followed the processes used within public health for conducting a critical literature review on a body of literature that includes a mixture of research-study designs. This operates in several stages which are outlined below. Given the volume of literature that had to be screened (12,000+ titles) the original objectives of the study were modified slightly.

1. Set criteria for which literature will be considered for inclusion (see below).
2. Conduct searches in the research-literature databases for health, psychology and anthropology.
3. Retrieve the papers restricted to that published in the English language or a language that can be translated by one of our existing staff as the cost of translation could not be accommodated in this proposal.
4. Read all the short-listed papers and:
   a. confirm that they meet the inclusion criteria
   b. determine the study design. Classify the artistic activity (see below)
   c. classify the health state (see below)
   d. classify the purported, and team determinations, of probable therapeutic mechanism.
   e. grade the quality of all the papers using the overall evidence hierarchy
5. Map the findings using a variety of classification systems and provide a narrative summary of the aggregate findings based on the consistency of the results across the various study designs, the overall quality of the evidence and the volume of available evidence.

Criteria for inclusion in the review

To be included in this review the publication must:

1. have been intended for a healthcare audience (practitioner, healthcare manager/provider or policy),
2. describe patient treatment/therapy
3. have a stated goal of providing evidence of therapeutic benefit from an artistic activity through measurement of a health state
4. the research study design must conform to one of the recognised research designs which includes experimental or quasi-experimental design, or observational designs (such as surveys or case-control designs). Qualitative designs will be recorded but excluded from this review. In addition, commentaries that only express an opinion (learned or not) will also be counted but excluded from the review as they are not considered to be any form of evidence
5. involve adults 18 years or over. Studies on children were excluded from this review for a number of reasons. It can be difficult to determine children's needs and some health conditions may not be established until later in life. Often research methodology has to be modified to accommodate the children's needs and circumstances, for example the involvement of parents or guardians. Instead, given the number of papers which had already been identified, it seemed preferable to recommend a further review of research on arts therapies with children and young people under the age of 18
6. focus on an artistic engagement that was primarily expressive/active rather than appreciative/passive.

Classifying the health state

The groupings that we found were determined by the quantitative research process; the inclusion of qualitative results would have produced a different classification. The literature reviewed was expected to contain an extensive array of 'health' measures. Classifying these for inclusion/exclusion generated lively debate among the team. Working together, we used our expertise to classify the measures we expected to see into a discrete set of categories for review. The starting categories are detailed below:

1) Standard/well recognised clinical measures of physical or mental disease
2) Generic or quality of life measures that incorporate a broad definition of 'health'. This is particularly relevant for our discussion on whether 'aging' was a health condition. We agreed that aging was a health condition, as much of the literature on health inequalities is premised on the recognition that health is socially graded: people in lower social and economic positions have an earlier onset of health problems, experience more severe disease, a greater number of co-morbidities and a shorter life expectancy than those higher than them in the social hierarchy.
Thus ameliorating the impact of aging by facilitating people in developing strategies to offset, manage or cope with their problems may either improve health or at least enable better quality of life experience, for example through the reduction of stress, distraction from pain, or encouraging the resolution of personal issues.

3) For patients who had limited abilities to interact with carers/researchers there will be measures of a patient’s behaviour. These are usually made by an observer rather than the therapist. An example is agitated behaviour in dementia sufferers.

4) After much debate, we excluded studies that only measured mood as we lacked the expertise in this complex and contentious area. We did, however, include such studies that looked at people with depression and bipolar disorder as these conditions have well recognised diagnostic criteria. The conceptual frameworks are vastly different between disciplines and this area is worthy of a review on its own.

Findings

Only a small number of quantitative studies using art therapy were found, although we recorded a significant body of qualitative research and learned opinion. The number of quantitative studies has increased hugely since 2011 (see Figure 3) and there is no reason to expect this trend not to continue. Just over half of the research was conducted in the US and UK with the remainder conducted in 14 other industrialised nations. The most frequently reported disciplines were psychology, psychiatry, music therapy and medicine. What was most disappointing was the dearth of co-authors from occupational therapy, rehabilitation or AT. On occasion, a therapist was described in the methods but did not appear to be a co-author.

Rather than reproduce the litany criticism that often accompanied previous critiques we elected to take a more constructive approach and highlight new review perspectives or solutions.

Figure 3 - Number of quantitative AT publications by year of publication
Classifying artistic activity

Classifying artistic activity is difficult as often interventions may comprise several elements: photography can often be combined with narrative, music listening is often used to stimulate writing or artwork. There are distinct disciplinary differences in how the art therapy is perceived to work, and these are reflected in the measures of 'health' and a poor choice of outcome measure can have a detrimental effect on the reported benefit of the art therapy. For example, the benefits of dance therapy are likely to be improved balance and co-ordination as well as improved social connections. Balance and co-ordination often take some time to improve, and short-term projects that focus exclusively on these elements may miss the more immediate benefit of friendship which improves adherence to the therapy. A small number of papers focused exclusively on physiological measures with a complete absence of any statement on the psychological or social benefits. A larger body of research acknowledges the social and psychological aspects but fails to use a measure that is fit for purpose. For example, art therapy in people with dementia incorporates cognitive function measures which are unlikely to detect any change over a short period of time. Overall many medically-based papers lack a broad conceptual framework about 'health'; however, the accepted quantitative studies tended to explore specific activities which could be reduced to 15 categories (see Table 1).

Studies tended to either give good definitions of the art medium employed for therapy (this was especially evident in papers that included a therapist as an author) or else poor ones, where perhaps the authors were more interested in the research process or medical prognosis than in the actual intervention. Many of the excluded studies involved passive use of art media. These studies frequently made assumptions that certain music forms were relaxing, although one of the rejected papers (Skelly and Haslerud, 1952) did attempt to evaluate their programme of music for passive listening before it was applied in the intervention. Often there was no attempt to investigate objectively the perception of music or literature before it was applied to the participants. While the results of experiencing Verdi, Bach or Persian poetry were sometimes meticulously measured, the authors had not considered whether the tonal range, rhythm or structure of the music or the reading might be significant; or where recordings were issued for individual use, whether factors such as auditory volume might be significant. There was often no mention of the context of the chosen art medium, i.e. the possibility that certain pieces may have specific associations both culturally and individually; although cultural acceptability was offered as a possible mechanism in an intervention for the elderly that used folk music (Eyigor 2009).

Almost none of the studies was able to determine how the effect change was produced by the intervention, and all concluded that more evidence would be needed to be certain of the cause-effect relationship.
Table 1 - Classifying the artistic activity in the abstracted publications

<table>
<thead>
<tr>
<th>Writing</th>
<th>Number of papers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expressive writing</strong>: an intervention (Pennebaker and Beal, 1986) usually consisting of three or four 20 minute sessions writing long-hand without pausing for correction on the deep emotional significance of a specific traumatic event</td>
<td>23 (includes 2 with active controls)</td>
</tr>
<tr>
<td><strong>Writing as therapy</strong>: the use of different forms of creative or autobiographical writing to express feelings or emotions or explore general topics in groups or classes</td>
<td>4</td>
</tr>
<tr>
<td><strong>Standard writing</strong>: a control writing task in which the participant is asked to write about the most significant emotional or traumatic event in their life (compared to a specific event in expressive writing)</td>
<td>1 (active control)</td>
</tr>
<tr>
<td><strong>Music</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Responding to music</strong>: moving, humming or singing along, clapping to music facilitated by experimenters or support workers, but excluding actual singing performance or playing instruments</td>
<td>8 (includes 3 with active controls)</td>
</tr>
<tr>
<td><strong>Music therapy</strong>: specific interventions designated as music therapy and facilitated by music therapists making music with participants</td>
<td>24 (includes 1 with active control)</td>
</tr>
<tr>
<td><strong>Music education</strong>: interventions which involve the teaching of an instrument such as piano keyboard</td>
<td>2</td>
</tr>
<tr>
<td><strong>Music performance</strong>: the involvement of participants in actual performance such as choir singing</td>
<td>6</td>
</tr>
<tr>
<td><strong>Art</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Art therapy</strong>: specific interventions designated as 'art therapy' and facilitated by art therapist.</td>
<td>13</td>
</tr>
<tr>
<td><strong>Arts based therapy</strong>: The use of arts media as a therapeutic intervention with a general purpose of activity engagement</td>
<td>16</td>
</tr>
<tr>
<td><strong>Dance</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Dance</strong>: the use of dance as a general activity where the emphasis is on performance or technique rather than exercise</td>
<td>2</td>
</tr>
<tr>
<td><strong>Dance therapy</strong>: Specific interventions designated as dance therapy and facilitated by dance therapists</td>
<td>2</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Storytelling</strong>: interventions involving the use of storytelling as an improvised narrative to talk about experiences</td>
<td>1</td>
</tr>
<tr>
<td><strong>Responding to reading</strong>: interventions involving the discussion of texts or literature.</td>
<td>2 (includes 1 with active control)</td>
</tr>
<tr>
<td><strong>Exercise</strong>: exercise used as a control task</td>
<td>2 with active control</td>
</tr>
<tr>
<td><strong>Dramatherapy</strong>: Specific interventions designated as dramatherapy and facilitated by therapists.</td>
<td>1</td>
</tr>
</tbody>
</table>
Measuring the Health

The literature reviewed did include an extensive array of ‘health’ measures which included physical and mental health, behaviours and social factors. The team found that it needed to move beyond the previous work on classification systems for health states (e.g. Rand Corporation) it was not always applicable to the populations included in the literature reviewed. Much of the literature was around mental health, cognitive function and stressful circumstances such as coping with a cancer diagnosis (see Table 2).

**Table 2 - Health conditions in the abstracted papers**

<table>
<thead>
<tr>
<th>Health Condition</th>
<th>Number of papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aging - generically about aging rather than cognitive decline</td>
<td>6</td>
</tr>
<tr>
<td>Arthritis or fibromyalgia</td>
<td>3</td>
</tr>
<tr>
<td>Cancer patients/survivors</td>
<td>11/2</td>
</tr>
<tr>
<td>Chronic conditions - living with long-term</td>
<td>4</td>
</tr>
<tr>
<td>Cognitive function</td>
<td>20</td>
</tr>
<tr>
<td>Unspecified mental health conditions</td>
<td>30</td>
</tr>
<tr>
<td>Neurological conditions</td>
<td>8</td>
</tr>
<tr>
<td>Quality of life</td>
<td>9</td>
</tr>
<tr>
<td>Stress</td>
<td>8</td>
</tr>
<tr>
<td>Sub-fertility</td>
<td>2</td>
</tr>
<tr>
<td>Surgical procedure</td>
<td>3</td>
</tr>
<tr>
<td>Well-being</td>
<td>5</td>
</tr>
<tr>
<td>Other - asthma, behaviour problems, emphysema or COPD (specifically about breathing), headache, musculoskeletal</td>
<td>22</td>
</tr>
</tbody>
</table>

The health conditions treated by art therapy are often those that conventional medicine struggles to deal with. As a result, the ‘health’ benefits are often difficult to measure, in particular quality of life or social connectedness. A significant portion of the literature reviewed here used measures of general well-being which can be regarded as ‘soft’ measures of health. It is also worth noting that global health literature may provide some alternative views of ‘health’ as there can be very different cultural assumptions contained in about the constituents of health and how a treatment works in non-Western treatment regimens. And, when working with people from other cultures a translation from western origins into non-Western languages there may be a need for cultural adaptation; for example, the context of Japanese or Korean society, or for that matter that of Korean migrants to the USA. For example, although Kim (2013) considers the effects of art making amongst South Korean seniors living in the USA, the context considered

---

1 the numbers add up to more than 92 as some conditions occur more than once
is that of the USA, not a specific South Korean cultural identity within it; nor is there consideration of whether the art making has particular significance to the participants, despite the mention of dual language use. Similarly, when considering a model of how social isolation links to depression, some account should be made of the contributing circumstances, particularly since the medium being employed are therapeutic arts and involve expression. Social isolation may often involve a trajectory, a historical process recounted in narrative specific to each person who experiences it. These issues may have a psychological effect on how life quality is perceived, but fall outside the measures of quality of life (QoL). This is generally conceptualised from the multi-domain WHO definition of health (“Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.”). Art therapy is also widely used in people with dementia where behavioural issues are the focus. A subset incorporated clinical measures of function such as a depression score or measure of activities of daily living but most failed to include behavioural measures. We acknowledge that the staff may find it difficult to measure behaviour, but appropriate checklists are common in other fields and the development of them is well understood, so we feel that there is a definite need for their inclusion in evaluations of AT.

Discussion

Art Therapy, like many areas of healthcare provision, suffers from a dearth of high quality evidence. An additional issue with therapeutic art is the underlying therapeutic mechanism. Proposed mechanisms are often driven primarily by the underlying conceptual framework of the authors. Hence researchers from disciplines such as medicine, psychology, folklore, cultural or anthropological knowledge will be motivated to adopt mechanisms they regard as suitable to the task in hand. This leads us to question whether any benefit to patients is derived from the interaction with the therapy provider, interaction with others, actual physical or neurological training or the learning of new practices and recognition of abilities. And we further question whether benefit derive from the use of specific approaches, personal goals set by the patient or is it simply that the ‘art’ is a the creative diversion for the people involved? As it is increasingly recognised that non-clinical (psychosocial) factors are key factors in health there has been an evolution in the models that we use to ‘explain’ how people become ill and how healthcare provision improves health. New models that incorporate social interaction with the therapy provider or other patients need to be examined more robustly and art therapy is particularly well placed to address this issue.

Despite this, the role and value of the arts in healthcare continues to be recognised and carry weight in policy terms. Although there is also evidence of silos of knowledge, Atkinson and White (2013) point out that there is a split between individualised notions of maintaining health and a more socially determined appreciation of the impacts of health inequalities and social interactions on everyone. Arts and health both operate across both sides of this divide but are rarely seen in this dual context.

In much of the research that did meet the inclusion criteria, the authors fail to suggest reasons why the arts therapy had an effect on the participants. For studies that found that arts therapy had positive benefits for the participants, it was often enough for the researchers that this was
the case. There was little attempt to probe how and why these results occurred. Conversely, studies that found that arts therapy did not have the expected benefits were more likely to offer underlying reasons; although in these cases study design was often cited as the reason (for example, see Rusted, 2006). This lack of a therapeutic mechanism is problematic on several fronts: a mechanism makes it easier to identify measures of effectiveness and assists in selecting a study design.

**Future Research/What next?/Recommendations**

Our findings in relation to evidence gaps are broadly consistent with arts reports (Arts Council 2014). The standard medical evidence hierarchy systems which classify research by study design are difficult to apply here for the reasons discussed previously. Alternative approaches – such as 'realist synthesis' might be used or other research synthesis techniques which are more inclusive. But there needs to be a balance of qualitative and quantitative research as the latter gives some measure of the burden on/benefit to society overall while the former provides information for healthcare service provision managers and planners.

Previous reviews of art therapy interventions have found a large number of observational studies which, because of their research design, cannot be used to draw conclusions about causality (i.e., we cannot say that treatment A improved condition B). This is a common occurrence within public health research as well and as a result causality criteria have been developed, with the set developed by Sir Bradford Hill being the most widely used. As a result a pragmatic approach is often taken such that, within public health, if 10 studies in 10 different populations using 3 different study designs with varying levels of quality and sources of bias all arrive at the same conclusion then we conclude that the proposed relationship between two factors is very likely to be genuine. The benefit is that research included in these reviews is also most likely to have been conducted in real-world situations, which suggests the treatment is both viable and practical. We have not seen Hill's criteria used in any of the reviews.

In addition, further qualitative work is needed to interview therapists and determine where the roadblocks are. We suspect that there is a real need for plain language, web-based information on how to conduct more robust evaluations of an Art Therapy programme and many other segments of healthcare. We need to move beyond the prescriptive "thou shalt" approach that predominates. For example, in our own practice we have found that many clinicians do not immediately understand the need for a comparison group - a fundamental component of a robust evaluation of treatment effectiveness. A future proposal will be made to develop more plain language descriptions of what is expected from a robust evaluation and specifically why these requirements are made. For example, why is blinding important or why is a comparison group needed?

**Collaboration**

Achieving adequate sample size is a problem for much arts therapy research. This can be down to various reasons: for example, the participant populations that tend to be used are notoriously difficult to recruit and conduct research upon (e.g. people with dementia), and some
forms of arts therapy may require one-on-one therapist-patient care (e.g. clay modelling). For these reasons, attaining reliable sample sizes for arts therapy research can be costly and time-consuming. It is telling that music therapy is the most prevalent form of arts therapy in our review, and this may be down to the fact that music therapy is often group-based and so can recruit larger numbers of participants.

One way of addressing this is for arts therapy researchers to engage in more collaborative work, across geographical areas and with researchers at other institutions. Although this happens already to an extent, greater collaboration would result in larger sample sizes and more reliable results.

**Identifying the underlying treatment mechanism**

Initially each paper was reviewed to determine what the authors felt was the purported mechanism by which the treatment is beneficial. This was often not explicit and had to be determined by a close reading of the text to assess framing or which literature is quoted to support a therapeutic effect, etc. The team used their cross-disciplinary expertise to develop an alternative model of how treatments may work. For example, is it the interaction with the therapy provider, interaction with others, or actual physical or neurological training? This group has the advantage of a principal investigator who has conducted research, over the past 25 years, on the non-clinical (psychosocial) factors that affect health. This perspective has the potential to radically alter our perspective on why therapeutic art is so effective. The team hopes to secure funding to develop this approach further but it is difficult to find funding for this research area.

**Areas of limited evidence/Evidence gaps**

There is little arts therapy research on dance, yoga and a range of activities which occupy a hinterland between art and exercise or art and crafts. These nuances have not been revealed in our focus on quantitative research. It was of concern how few occupational therapists were involved in quantitative research. In general, most of the research was not led by practitioners and in some papers, although they were clearly involved in the delivery of the interventions, they were not named as authors. In many journals the discipline of the authors or their qualifications is not listed. This may be a problem of professional recognition and visibility.

**Theoretical Understanding**

Research in this area needs to do more to suggest reasons why arts therapy interventions are effective. Is it because of engagement with the art form itself? Is it a result of being part of a group? Is it because of involvement with a research project? Perhaps the answers to these will be speculative, but more needs to be done on this front. Suggestions for ways forward include:

- arts therapy researchers working more closely with researchers from ‘purely’ arts disciplines (e.g. English, Art History, Fine Art departments) in order to establish what effect arts engagement might be having
- arts therapy researchers conducting follow-up studies as routine practice in order to establish longer term effect of arts interventions
- AT researchers need to work with research methodologists to develop stronger designs. This does not necessarily mean conducting only RCTs but it does mean more rigorous design.
References and external links

References (in report)


References that were abstracted


Appendix

Search strategy

The following databases were searched:

- ASSIA (ProQuest)
- The Cochrane Library (Wiley)
- MEDLINE (EBSCO)
- PsycINFO (ProQuest)
- Scopus (Elsevier)
- Sociological Abstracts (ProQuest)
- Web of Science (Thomson Reuters).

All searches have been written up for MEDLINE using the EBSCO interface.

Explanation of search terms used: / = MeSH Heading; exp = exploded MeSH Heading; asterisk (*) denotes any character; N = adjacency of words; "" = phrase search ti = title word; ab = abstract word; all = all fields; so = journal name

1. health.ti,ab
2. healthcare.ti,ab
3. "health care".ti,ab
4. "clinical outcome*".ti,ab
5. wellbeing.ti,ab
6. well-being.ti,ab
7. "well being".ti,ab
8. wellness.ti,ab
9. therapeutic*.ti,ab
10. exp health/
11. delivery of health care/
12. outcome assessment (health care)/
13. therapeutics/
14. or/1-13
15. "therapeutic art*".ti,ab
16. "art therap*".ti,ab
17. writ* N2 story*.ti,ab
18. writ* N2 stories.ti,ab
19. writ* N2 play*.ti,ab
20. writ* N2 poetry.ti,ab
21. writ* N2 poem*.ti,ab
22. writ* N2 movie*.ti,ab
23. writ* N2 film*.ti,ab
24. writ* N2 music*.ti,ab
25. produc* N2 play*.ti,ab
26. produc* N2 video*.ti,ab
27. produc* N2 film*.ti,ab
28. produc* N2 movie*.ti,ab
29. direct* N2 play*.ti,ab
30. direct* N2 film*.ti,ab
31. direct* N2 video*.ti,ab
32. direct* N2 movie*.ti,ab
33. mak* N2 movie*.ti,ab
34. mak* N2 video*.ti,ab
35. mak* N2 film*.ti,ab
36. creat* N2 video*.ti,ab
37. creat* N2 movie*.ti,ab
38. play* N2 music*.ti,ab
39. play* N2 instrument*.ti,ab
40. perform* N2 dance*.ti,ab
41. perform* N2 music*.ti,ab
42. perform* N2 play*.ti,ab
43. perform* N2 drama*.ti,ab
44. "computer animation*".ti,ab
45. sewing.ti,ab
46. singing.ti,ab
47. sculpture*.ti,ab
48. embroider*.ti,ab
49. quilting.ti,ab
50. crochet*.ti,ab
51. knitting.ti,ab
52. “fibre art*”.ti,ab
53. “fiber art*”.ti,ab
54. “wood work*”.ti,ab
55. woodwork*.ti,ab
56. paint* N5 art.ti,ab
57. paint* N5 artwork*.ti,ab
58. draw* N5 art.ti,ab
59. draw* N5 artwork*.ti,ab
60. draw* N5 pictur*.ti,ab
61. photography N5 art*.ti,ab
62. photography N5 pictur*.ti,ab
63. mak* N/2 artifact*.ti, ab
64. mak* N/2 artefact*.ti,ab
65. creat* N/2 artifact*.ti,ab
66. creat* N/2 artefact*.ti,ab
67. art* N/2 installation*.ti,ab
68. perform* N/2 poetry.ti,ab
69. publish* N/2 poetry.ti,ab
70. publish* N/2 story*.ti,ab
71. publish* N/2 stories.ti,ab
72. writ* N/2 narrative*.ti,ab
73. writ* N/2 autobiograph*.ti,ab
74. read* N/2 play*.ti,ab
75. bibliotherapy.ti,ab
76. drum* AND music*.ti,ab
77. "serious play".ti,ab
78. creativ* N/2 art.ti,ab
79. creativ* N/2 arts.ti,ab
80. storytelling.ti,ab
81. story-telling.ti,ab
82. "story telling".ti,ab
83. photovoice.ti,ab
84. "men's shed"*.ti,ab
85. "oral histor*" and narrative*.ti,ab
86. "oral histor*" and memor*.ti,ab
87. "oral histor*" and interview*.ti,ab
88. art therapy/
89. bibliotherapy/
90. dance therapy/
91. music therapy/
92. sensory art therapies/
93. or/15-92

94. mouse.all
95. mice.all
96. hamster*.all
97. bird*.all
98. biofilm*.all
99. "polymer film"*.all
100. enzyme*.all
101. protein*.all
102. medical problems of performing artists.so
103. or 94-102

104. 14 and 93
105. 104 and not 103
106. 105 English Language only
### Appendix Table 1 - Search numbers yielded by database

<table>
<thead>
<tr>
<th>Database</th>
<th>Search numbers yielded</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASSIA (ProQuest)</td>
<td>639</td>
</tr>
<tr>
<td>The Cochrane Library (Wiley)</td>
<td>1009</td>
</tr>
<tr>
<td>MEDLINE (EBSCO)</td>
<td>2224</td>
</tr>
<tr>
<td>PsycINFO (ProQuest)</td>
<td>5174</td>
</tr>
<tr>
<td>Scopus (Elsevier)</td>
<td>3596</td>
</tr>
<tr>
<td>Sociological Abstracts (ProQuest)</td>
<td>412</td>
</tr>
<tr>
<td>Web of Science (Thomson Reuters)</td>
<td>4007</td>
</tr>
</tbody>
</table>

**Total number of results**  
17061

**Total number of results with duplicates removed**  
12122