UNIVERSITY OF BIRMINGHAM

THE USE OF DRAMA BASED THERAPY WITH FORENSIC POPULATIONS

by

Lucy Brierley

A thesis submitted to the

University of Birmingham

For the degree of
Doctorate in Forensic Psychology (FOREN.PSY.D)

Centre for Applied Psychology

School of Psychology

College of Life and Environmental Sciences

August 2021

ABSTRACT

The focus of the present thesis is the use of Drama based Therapy (DbT) with forensic populations. Research has suggested that DbT can have a positive impact on vulnerable individuals in a range of settings. DbT has been used in forensic settings such as prisons alongside standard manualised interventions or, in some cases, as an alternative to these interventions. Chapter 1 presents an overview of the use of Arts therapies in general and, more specifically, the use of DbT with forensic populations. In addition, an overview of the aims of each chapter of the thesis is provided. Chapter 2 presents a systematic review of the existing literature exploring the use of DbT with forensic populations. Findings from 10 papers included in the review revealed that there is a range of DbT methods being used, and that the DbTs lead to positive outcomes such as improved social and communication skills, reductions in anger, improvements in mental well-being, and an increase in hope for the future. Findings are discussed with reference to implications for practice, and recommendations for future research are made. In order to address one such recommendation, Chapter 3 explores the views and experiences of Geese Theatre Company (GTC) Actor Practitioners (APs) who deliver DbT to individuals in forensic settings. Thematic analysis was used to draw themes from the data obtained through semi-structured interviews with current and former APs. Participants spoke about the positive impact that the GTC intervention had on those who engaged with it. They also spoke of the qualities and skills that they felt were necessary for the job role, the benefits of the flexible and creative nature of the DbT, and lastly, of the challenges they experienced in their role. Chapter 4 presents a critique of the Warwick Edinburgh Mental Well-being Scale (WEMWBS; Stewart-Brown et al., 2007). Findings revealed good psychometric properties of the WEMWBS, such as construct and content validity, and test-retest reliability. However, it is suggested that further research is required to better establish the reliability and validity of the scale. Chapter 5 presents a summary of the thesis' main findings and provides some key

limitations and recommendations for future research. In addition, implications for practice are provided with the hope that these will be of interest to professionals who use DbT with forensic populations. It is also hoped that, more broadly, the findings could add to the evidence base regarding *what works* in rehabilitative interventions.

Acknowledgments

I would like to massively thank my academic supervisor - Dr Zoe Stephenson, for her continued support and encouragement. Your advice and patience have been greatly appreciated throughout this process. Thank you, Nicola Schaum, – for all the proofreading you've done.

To my Dad (my everything!), your continued love and support means so much. You are every kind of support to me.

To my niece and nephew, for making my life meaningful.

Table of Contents

List of Tables

List of Figures

List of Appendices

CHAPTER 1

Introduction

Since the beginning of the 1900's, there has been a shift from the punishment of offenders towards rehabilitative efforts in the criminal justice system (Cullen & Gendreau, 2000). Martinson (1974, as cited in Sarre, 2001) assessed over 200 studies which looked at prison rehabilitative programmes. Martinson (1974) concluded that the majority of psychological treatment programmes failed to reduce recidivism. Nonetheless, rehabilitative efforts have continued and there are many evidenced based, accredited psychological interventions used in forensic settings worldwide (Cullen & Gendreau, 2000; 2001). Whilst methodological limitations (e.g., a lack of a comparison group) make it difficult to ascertain whether or not a programme *works* (Sapouna et al., 2015), Cullen and Gendreau (2000) suggested that "our 'best bet' for reducing recidivism and improving the lives of those processed through the correctional system is to involve them in rehabilitation programs that have therapeutic integrity" (p. 161).

Based on the findings of research in the field of offender rehabilitation, suggestions have been made regarding the principles that programmes should adhere to in order to be effective. For example, Andrews and Bonta (2010) state that programmes used in UK community settings, which adhere to the Risk Need Responsivity (RNR) principles reduce criminal recidivism. Whilst accredited programmes have shown to benefit some offenders, it is clear that they do not have a positive impact on all those who take part, in particular those with poor literacy skills (Andrews & Bonta, 2010). Consequently, alternative approaches are needed in the criminal justice system to better address the needs of offenders (which include people with developmental and learning difficulties, mental health difficulties, and people with personality disorders). In respect of the responsivity principle, it is suggested that individual differences in offenders and different learning styles are taken into consideration when designing and delivering interventions (Andrews & Bonta, 2010).

Arts Therapies

It has been suggested that the arts (e.g., music, painting, pottery, dance, and drama) and arts based interventions offer an alternative method of therapy which some may find more beneficial than manualised, accredited interventions. Arts based interventions have been used in the criminal justice system for many years with a range of aims, for example, "to heal, educate, 'reform', or to improve self-esteem, emotional literacy and aid socialisation by providing creative opportunities for self-understanding and expression" (Meekums & Daniel, 2011, p. 229). In their systematic synthesis of the literature, Meekums and Daniel (2011) concluded that despite methodological limitations, arts therapies used in criminal justice settings can have "a positive, humanising and healing effect" (p. 223).

More specifically, and with reference to the focus of this thesis, one principle of good practice in offender rehabilitation is that "learning should be active, concrete and based on the use of role play to practise new skills" (Baim et al., 2002, p. 19). Therefore, exercises such as role playing are used within some offender behaviour programmes (OBPs) and drama based therapies (DbTs) are now widely recognised as a method of psychological intervention in UK prison and forensic mental health settings (Baim, 2007).

Drama Therapy

Drama therapy is used both as a banner term to refer to a range of drama based therapies (Gersie, 1996), and as a stand-alone term to refer to a specific form of drama therapy. With reference to the latter, a Drama Therapist uses dramatic performance as a therapeutic tool to address mental wellbeing (Pendzik, 2006). Drama therapy has been defined as "…an active, experiential approach to facilitating change" (North American Drama Therapy Association, 2014, paragraph 1) and as a type of psychological intervention where performance arts are used to encourage the therapeutic relationship (The British Association of Drama Therapists, 2011). With reference to drama therapy as a banner term, it

has been noted that drama therapy is interdisciplinary in nature in that techniques/approaches used under this banner are rooted across disciplines such as anthropology, psychotherapy, philosophy, and theatre (Landy, 1994). It is acknowledged that there are a range of methods associated with drama and theatre which aim to facilitate change in an individual (Gersie, 1996). Each drama based approach may have a unique and different set of components, methodology, and therapeutic process. Despite variations, however, there are substantial similarities across a range of modalities of drama based therapies (Gersie, 1996). Jones (2007, as cited in Lištiaková, 2015) suggested that "embodiment or dramatization of body, projection, empathy and distancing, role play and personification and the connection of drama and everyday life and its transformation" (p. 21) are common core processes across drama/theatre approaches. Furthermore, Valenta (1995 as cited in Pendzik, 2006) reported that reflection is a common component across drama therapy approaches. Reflection is explained as a mirroring process in drama to promote therapeutic change. It involves participants gaining an understanding of their emotions and experiences through watching metaphorical representations of people and scenarios.

One approach which comes under the banner of drama therapy is psychodrama (Gersie, 1996; Lištiaková, 2015). Psychodrama involves role training with an end goal of "…achieving catharsis, insight and transpersonal spiritual connection" (Lištiaková, 2015, p. 22). In short, it seeks to address health and wellbeing over a period of time through the use of drama techniques. An aim of psychodrama is to help an individual gain control over their emotions (Baim et al., 2002). The British Psychodrama Association (2020) state that "issues or problems and their possible solutions are enacted rather than just talked about. Psychodrama offers the opportunity to practise new roles safely, see oneself from outside, gain insight and change" (British Psychodrama Association, paragraph 3). Drama therapy and psychodrama share common factors, however, drama therapy veers more towards "creative-expressive learning of

roles whereas psychodrama is oriented more toward experiential learning, including specific working through of emotional, cognitive, interpersonal, behavioural and non-specific issues" (Kedem-Tahar & Felix-Kellermann, 1996, p. 34).

Other methods/techniques which come under the banner term of drama therapy include, for example, role theory (Landy, 1994) and applied theatre (Baim, 2007). Landy (1994) proposed that life can be a performance and that each individual plays a set of dynamic roles throughout their lifespan. He proposed that through practicing a range of roles, an individual will be better able to cope with potential challenging life situations. Applied theatre practices use theatre as a shared and safe platform for the expression of stories, role plays, and simulations. It draws from components of theatre performance to guide theatre experiences for participants whereby "the boundaries between actors and spectators are purposefully blurred as all participants are involved as active theatre makers" (O'Connor & O'Connor, 2009, p. 471). Applied theatre practices have been used in order to create social change, such as prison theatre, in an attempt to transform such communities (O'Connor & O'Connor, 2009).

As mentioned above, there is much common ground across a range of drama therapy approaches. For the purpose of this thesis, the term 'drama based therapies' (DbTs) will be used throughout when referring to therapies which use drama based techniques to address the needs of a group of individuals. Information pertaining to the particular approach or techniques used in specific DbTs are summarised in Chapters 2 and 3.

The Impact of Drama based Therapies (DbTs) within Forensic Populations

Globally, DbTs are increasingly used with offender populations (Fazel & Seewald, 2015). Multiple studies have shown such programmes to have a range of positive effects on individuals in the criminal justice system in the UK (see Chapter 2 for a review of the literature). For example, Stallone (1993) investigated the impact of a psychodrama intervention on criminal recidivism and found that following the intervention, participants had a more

objective view of their life in prison, and their antisocial behaviour reduced. Other researchers have found that DbTs can reduce stress (Haskell, 1960, as cited in Harkins et al., 2011), help participants feel less isolated (Matarasso, 1997), and be more cooperative and trusting of others (Phillips, 1997). Furthermore, studies have found DbTs to result in improvements in self-esteem and self-worth (Phillips, 1997) as well as increased levels of confidence in approaching and communicating with people close to them (Melnick, 1984). A study investigating the benefits of DbTs with young offenders found that it improved perspective taking and reduced rigidity in thinking (Chandler, 1973).

As previously mentioned, methodological difficulties when evaluating interventions are also common in studies looking at the impact of DbTs. For example, there is often a lack of a control group, no follow-up measures (i.e., to explore whether any positive impact was long lasting), and small sample sizes meaning that results cannot be generalised. In addition, participants may be taking part in multiple interventions in prison settings, so it is not possible to firmly conclude that any improvements are as a direct result of engaging in DbT. However, despite methodological limitations, existing findings are promising in showing a trend towards DbTs having a positive impact on individuals in forensic settings in both the UK and US.

High levels of learning and developmental difficulties have been recorded in forensic populations in the UK (Antonowicz & Ross, 1994; Chesner, 1995) as well as issues such as impulsivity, lack of adaptive social skills, low self-esteem and confidence (Ross & Hilborn, 2008). These factors may pose barriers to engagement in more traditional Offender Behaviour Programs (OBPs). DbTs are thought to be more accessible for forensic populations in that participants do not have to meet a certain level of literacy in order to take part, and the delivery can be tailored to the needs of the group, something which may be more difficult to achieve in a manualised OBP. Hughes (2005) highlighted a need to focus on engaging forensic

populations on an imaginative level to address well-being; an aim which can be met through DbT.

Aims of the Thesis

Broadly speaking, this thesis aims to add to the knowledge base around the use of drama based therapies (DbTs) within forensic populations. More specifically, it aims to explore research on the impact of DbTs within forensic populations and also aims to provide drama practitioner perspectives of a specific type of DbT. Despite the increasing use of DbTs in forensic settings over the last few decades, there remains a dearth of evidence regarding the implementation and effectiveness of such interventions. As such, it is hoped that the following systematic literature review and empirical research project will provide further insight into this topic. In addition, with a view to evaluating DbTs, there is a need to investigate tools used for this purpose. To address the above needs, this thesis presents the following chapters:

- Chapter 2 presents a systematic literature review exploring the use of drama based therapy (DbT) within forensic populations. The review uses a narrative data synthesis model to present the findings of qualitative, quantitative and mixed method design studies. Articles were screened against inclusion and exclusion criteria and were assessed for methodological quality. The findings are discussed in relation to implications for practice, and recommendations for future research opportunities are made.
- Chapter 3 is an empirical study exploring Actor Practitioner's (APs) views and experiences of the use of applied theatre (a form of DbT) within forensic populations. Participants were sourced from the Geese Theatre Company (GTC), a UK based non-profit organisation. More specifically, the study aimed to gain insight into: aspects of the programmes APs felt were beneficial to clients; techniques they used to engage clients; the therapeutic relationship; training they have received; as well as challenges they have

experienced within their role as Actor Practitioners. Data were collected using semi-structured interviews, and Thematic Analysis was employed to analyse the findings. Key themes and sub-themes were collated and discussed in relation to existing literature and suggestions relating to practice and potential future research avenues are made.

- Chapter 4 provides a critique of The Warwick-Edinburgh Mental Well-Being Scale (Tennant et al., 2006). This inventory was selected for critique as it has been used to evaluate the efficacy of a DbT delivered by the Geese Theatre Company. The reliability and validity of the measure are discussed in relation to the inventory's utility for practice.
- Chapter 5 provides a summary of chapters 2, 3, and 4 and findings are discussed with reference to the wider literature base. Lastly, implications of practice are evaluated more broadly and suggestions for future research projects pertaining to the use of DbTs are made.

CHAPTER TWO

The use of Drama Based Therapy (DbT) with Forensic Populations: A Systematic Review of the Literature

Abstract

Drama based therapy (DbT) has long been used as a therapeutic tool with vulnerable populations. However, there is a lack of research regarding the use of DbT with forensic and mental health client groups found in secure settings (e.g., secure hospitals). The aim of this systematic literature review was to explore and synthesise research regarding the use of DbT with individuals in forensic settings, including those with mental health issues. The review will provide detail regarding the types of DbT being used in forensic settings, and then synthesise findings of qualitative and quantitative studies which explored the impact of DbT.

A systematic search of the following electronic databases was conducted to include research published from 1990 to 2020: National Criminal Justice Reference Service (NCJRS); Criminal Justice Database; ESSCO; Ovid; ProQuest; Web of Science; Humanities & Social Sciences Collection; International Index to Performing Arts; PubMED; and PsycINFO. Articles were screened against inclusion and exclusion criteria and were quality assessed using the CASP and AXIS tools. A narrative data synthesis model was used.

Ten articles (four qualitative, four quantitative, two mixed methods) met the inclusion criteria and were considered of sufficient quality to be included in the review. Findings are interpreted with caution given the methodological limitations of the studies included in the review. However, the reported positive impact that DbTs have had on issues such as anger, aggression, and mental wellbeing within this population is promising. Limitations and strengths of the articles are discussed, and recommendations are made for future research.

Introduction

A wide range of rehabilitative interventions are used with forensic populations in prisons, secure units and the community in the UK and other developed countries. Often these are CBT based, accredited, and manualised programmes which aim to address maladaptive thoughts and behaviours and take place on a one-to-one basis or, more commonly, in a group setting. More specifically, such interventions address criminal behaviour and developing skills such as problem solving, emotional regulation, as well developing and maintaining relationships with others (Bergman, 2017).

Many interventions undergo internal and external evaluation to ascertain the extent to which they are effective in reducing levels of recidivism (Bergman, 2017). Despite findings suggesting that engaging in such interventions can be effective in some cases, re-offending rates of those who complete interventions indicate that the interventions do not *work* for all individuals (Lowenkamp et al., 2012).

Offender populations typically present with a range of issues including, for example, poor emotional control, low self-esteem and/or confidence, difficulties in maintaining relationships, impulsivity, and pro-criminal attitudes (Andrews & Bonta, 2010). The Risk, Needs, Responsivity principles (RNR) propose that decisions regarding which programme an offender should complete should be based on an individual's level of risk, their criminogenic needs, and in reference to the responsivity principle, that it should also be tailored to the individual (i.e., taking individual level factors such as learning difficulties and level of motivation into account). One criticism of some current manualised interventions is that they take a *one-size-fits-all* approach and therefore the content and delivery is not suited to a wider cohort of offenders (Lowenkamp et al., 2012). Hence, it is suggested that there is a need for more research into the types of therapeutic work that may benefit forensic populations, emphasising the need to veer away from the one-size-fits-all approach by using a range of

techniques to match the needs and learning styles of participants.

As outlined in Chapter 1, arts based therapies (e.g., art, drama, music, and dance) have been shown to have promising outcomes for offender populations (Meekums & Daniel, 2011). More specifically, Drama based Therapy (DbT) (See Chapter 1) is increasingly being used with forensic populations, showing promising results (Feniger-Schaal & Orkibi, 2020; Jennings et al., 1997). Drama based Therapy has been described as "…an active and experiential psychotherapy modality that involves the intentional and systematic use of drama/theatre processes as primary means to achieve psychological growth and change within a psychotherapeutic relationship" (Feniger-Schaal & Orkibi, 2020, p. 1).

It has been suggested that, through DbTs, offenders can acquire new perspectives and insight or understanding as to the factors that led to their conviction or hospital admission. Furthermore, DbT can result in clients gaining insight into: the barriers that hinder them in changing negative behavioural patterns; how behavioural patterns repeat over time; and how maladaptive behaviours can be learned and modelled from a young age (Feniger-Schaal & Orkibi, 2020). In a study by Harkins et al. (2011) exploring the impact of a DbT intervention with an offender population, one male participant commented:

> It [the DbT] gave me confidence. I never had confidence before. It gave me strength and made me realize a lot of things. It made me think deep and open up my eyes. . . This is the first time I've had this experience in my whole life. It was perfect, more than perfect. It made me open my eyes, wake up and smell the coffee. . . It made me happy and proud. It was an amazing job they've done. You don't understand how much I appreciate it. It has so much meaning to me. Brilliant. Brilliant. (p. 553).

Arts based therapies such as DbT are considered to be more effective for some prisoners than some manualised programmes used in prisons and secure units in the UK and

other countries (Baim et al., 2002). In large part, this is thought to be due to the ability of DbTs to be tailored to the needs of an individual (see Chapter 1) by not requiring a certain literacy level and tapping into an individual's emotions and feelings in a way that manualised programmes may struggle to do (Baim et al., 2002). In addition, through acting, DbT provides clients with the opportunity to practice skills (e.g., social skills and self-control when placed in a challenging situation) which they may later need to draw upon whilst incarcerated or following release. Further to this, DbTs often place an emphasis on supporting clients in self-reflection (e.g., reflecting on their past, their thoughts and behaviours, and their future goals). The use of role-play and elements of the Cognitive Behavioural Therapy approach have been found effective for many programmes used in the criminal justice system (Antonowicz & Ross, 1994; Blacker, 2008); these techniques are key in DbT, which may explain, in part, why the findings of evaluations of DbT programmes are promising.

Various research designs have been used to measure the impact of DbTs in forensic populations (Meekums & Daniel, 2011). However, it is important to note that there remain methodological difficulties in measuring the impact of DbTs in forensic populations. For example, often the studies are qualitative, and where quantitative studies exist, they lack a control group. Nonetheless, there is an emerging body of research, using both quantitative and qualitative methods, which has explored the impact of a range of DbTs and will be considered in this review.

Rationale for the Review

Evidence suggests that DbTs can have a positive impact on patients in mental health settings (Feniger-Schaal & Orkibi, 2020; Meekums & Daniel, 2011), however, less is known about the use of DbT with forensic populations. As outlined in Chapter 1, DbTs vary in content and delivery (i.e., with some falling under the definition of psychodrama or applied

theatre rather than drama therapy), and as such, the findings of one study cannot be generalised to all types of DbTs. However, there are commonalities amongst the range of DbTs available in forensic settings (See Chapter 1 for further detail) and it is considered beneficial to collate this research to gain an understanding as to what types are being used and what impact they are having. As previously mentioned, there are methodological difficulties in evaluating DbT interventions (See Chapter 1 for further detail). In addition to quantitative studies, the review will include qualitative studies in order to gain greater insight into the impact that DbTs are having on participants within forensic settings.

Aims of the Review

Broadly speaking, this review aims to systematically examine peer-reviewed literature regarding the impact of DbT in forensic populations. More specifically, the review will synthesise findings from both qualitative and quantitative literature in respect of DbT outcomes (e.g., changes in scores on self-report measures, comments/feedback provided by participants). In addition, in order to comment on commonalities between DbTs, a description of each DbT will be provided. In short, the aims of the current review are:

- *To explore the types of DbT interventions used with forensic populations.*
- *To synthesise literature regarding the impact of DbT interventions with forensic populations.*

Method

Scoping exercise

A scoping search using Cochrane's Database of Systematic Reviews (CDSR) was conducted to ascertain whether there were any existing literature reviews exploring DbTs with forensic populations. In addition, further relevant databases were searched for the

presence of literature reviews on this topic: PsycINFO, National Criminal Justice Reference Service, Web of Science, and the internet search engines Google and Google Scholar. The CDSR search yielded one article by Meekums and Daniel (2011). This article described a literature synthesis looking at arts therapies with offender populations. This review, however, included all creative arts therapies (e.g., music, dance, and illustration) rather than placing a specific focus on DbT and was conducted nearly a decade ago. The scoping search also indicated that a number of relevant studies were published in the last decade. As such, it was concluded that an updated systematic literature review on the topic of DbT with forensic populations was warranted.

Search strategy

Identifying key words and terms

From the few studies found during the scoping task, a search strategy was formulated. The systematic literature review (Meekums & Daniel, 2011) identified during the scoping search was investigated. Alongside this, the author of this review spent time developing relevant search terms/key words based on the study's aims and objectives in order to capture relevant literature. Additional search terms were highlighted from the few articles found through the scoping exercise (e.g., psychodrama). All key words and terms were inputted into the PsycINFO database one at a time, initially looking at key words and terms in titles and abstracts of articles. Truncations were used for key words and terms. All terms were combined using AND or OR as shown below:

Prison* OR Offend* OR Inmate* OR Crim*

AND

Drama* OR Theatre* OR Creative* OR Art* OR psychodrama (see Appendix K for combinations).

Identifying additional key words/terms within the topic area was an iterative process. It was recognised that much of the DbT research was hidden under the wider term *Art Therapy.*

Development of the Inclusion Criteria

The decision was taken to focus on studies over the last two decades. It is noted that processes and approaches in the use of DbT have evolved over the years and as such it was noted that papers published from the year 2000 would represent a current reflection of DbT in forensic and mental health settings. Through setting this limit it was possible to glean a more homogenous set of articles. Additional criteria for inclusion are presented in Table 1.

Table 1

Inclusion and exclusion criteria

	Inclusion	**Exclusion**
Sample	Studies where all clients are within a forensic populations (including juvenile and adult samples)	Studies that include clients with no forensic link
Phenomenon of Interest	Studies that describe a DbT intervention, e.g., drama therapy, psychodrama, applied theatre	Studies that describe other arts based therapy interventions (e.g., music, dance)
Design	Questionnaire; survey; interviews; focus groups; pre/post psychometric tests.	No empirical data collection method used

Evaluation	Themes; views; experiences; attitudes; opinions; perspectives *OR* Outcome measure from psychometric test or questionnaire.	No empirical data analysis method used
Research Type	Qualitative, quantitative, and mixed methods.	Non-empirical papers
Other	Studies published in the English language. Studies published from 2000 to present day. Studies that show completeness in terms of their write-up. Peer reviewed, published articles.	Books, Non-empirical papers, Systematic Literature Reviews, Meta-analyses.

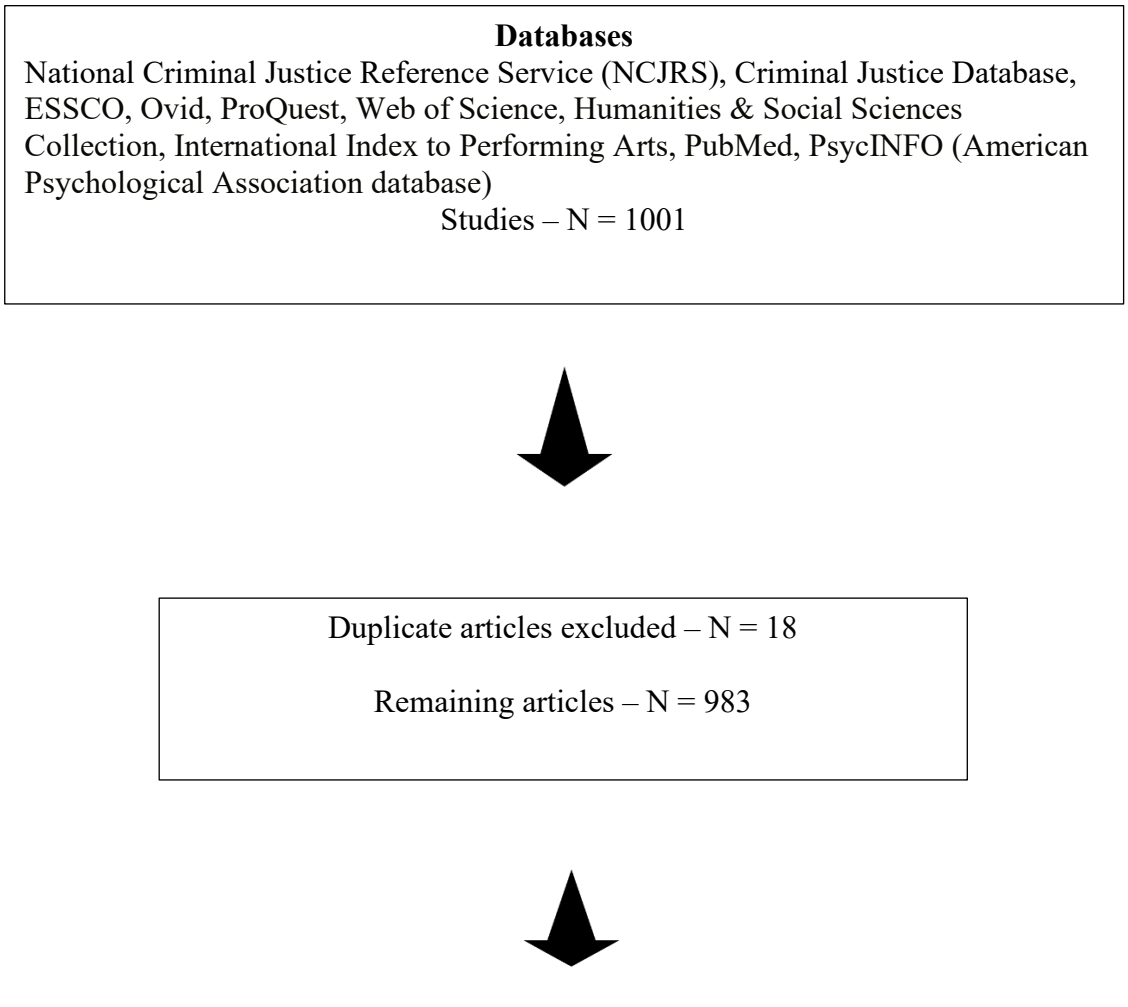
Sources of literature

Databases were selected based on how relevant they were to the topic area. The following electronic databases were searched: National Criminal Justice Reference Service (NCJRS); Criminal Justice Database; ESSCO; Ovid; ProQuest; Web of Science; Humanities & Social Sciences Collection; International Index to Performing Arts; PubMED; and PsycINFO (American Psychological Association database). All databases were searched from

the year 2000 to 2020. References were structured through the EndNote software. This allowed the author to scan for duplicates, and structure the results of searches by using file names. A further search stage took place whereby reference lists of the articles that passed the inclusion criteria (N=10) were searched for any additional relevant articles.

Summary of Search Strategy

A total of 1001 publications were identified from the database searches. Eighteen duplicates were removed leaving a total of 983 publications. Articles that were deemed to be unrelated based on the title and/or abstract were discarded. Thirteen articles remained, and an additional two articles were identified from a manual search of reference lists. These fifteen articles were read in full and screened against the inclusion criteria which resulted in the final inclusion of ten studies.



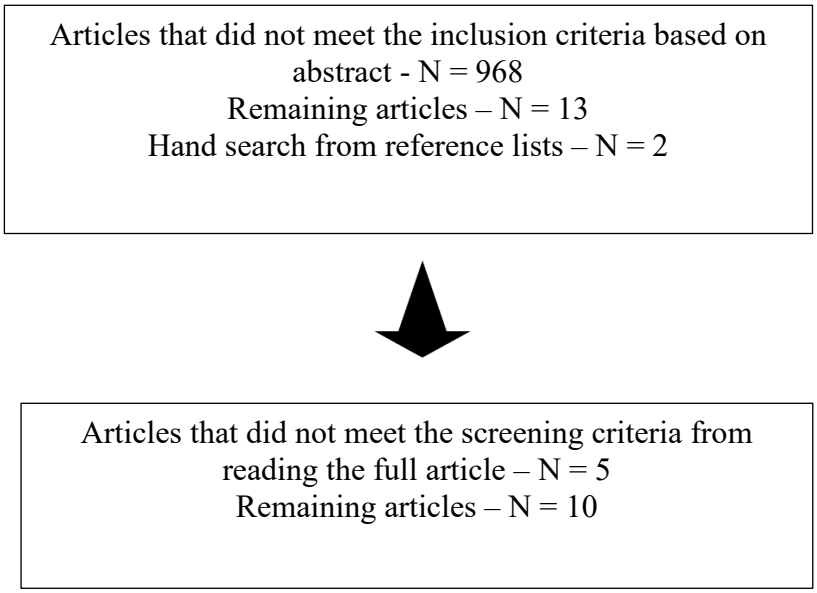


Figure 1: Flow chart of the search strategy process

Quality Assessment

The ten studies selected for inclusion in the review (four quantitative, four qualitative, two mixed-method studies) were then assessed for methodological quality to ensure that articles included in the review were of an acceptable standard. Studies varied in the design used, and as such, it was deemed necessary to use two quality assessment tools.

Qualitative studies

The Critical Appraisal Skills Programme (CASP, 2010) is a measure used to appraise research papers. Versions of the CASP checklists are available for a range of study designs, but for this review the CASP Qualitative Checklist (Critical Appraisal Skills Programme, 2010) was used to assess the qualitative studies and qualitative elements within mixed design studies for inclusion in the review (n = 6). In short, this appraisal tool aims to assess: the validity of a study; the findings/outcomes of the study; and how meaningful these outcomes are in the study's research area i.e., the generalisability of the findings.

The CASP tool comprises initial screening questions, followed by more detailed questions regarding the individual sections of an article. Adaptations made by the author included: adding *somewhat* to the dichotomous rating scale. This allowed for other options to be ticked should the information in the article be unclear and not falling into a yes/no response. In addition, the following item was added: "Did the researcher consider their own potential impact on data collection, data analysis, and their findings" (Appendix B). The rationale for adding this item was minimise researcher bias.

Cross-sectional Studies

The AXIS tool (Downes et al., 2016) was selected to evaluate the quality of quantitative elements of the mixed method studies and the quantitative studies (n = 6) as they were of cross-sectional design. This tool was chosen as it assesses a study's design as well as the way findings are analysed and reported in articles. Slight adaptations were applied to this tool (Appendix A). The author assessed the quality of all studies, and a quality assessment score was allocated (Appendix C). As the scale does not provide a numerical outcome value this was applied as it was considered helpful for rating the papers.

Inter-rater Reliability Agreement

To address rigour of the quality assessment carried out, the author of this review contacted a colleague from the University of Birmingham to review three articles (i.e., 33% of the total papers; two qualitative papers and one quantitative paper) using the adapted quality review forms. There was 95% agreement in terms of scores/ratings which was considered sufficient (Armstrong et al., 1997). An agreement of 70% or more was deemed acceptable based on a search of inter-rater reliability papers (Belur et al., 2021). All studies were deemed to be of acceptable quality and were therefore all included in the review.

Data Extraction

A data extraction form (Fleeman & Dundar, 2014) was adapted to meet the aims and objectives of this review (Appendix D). The form incorporated information such as study title, author/s, and when it was published. Study aims and objectives, the design of the study, measure/s used, basic information on the sample used, data analysis, study findings, and study key limitations were also included in the form. In addition, in line with the aims of this review, it was considered necessary to extract information from each article regarding the type of DbT that was used in each study (Table 3).

Data synthesis

The included studies varied in terms of assessment measures used for the quantitative studies; research designs; research questions; and methods of data analysis. Due to this lack of homogeneity, a narrative synthesis was decided upon as the method for analysing the data in this review (Pope et al., 2007). Narrative synthesis can be used for synthesising both qualitative and quantitative data by describing methodology and study findings (Popay et al., 2006).

A framework for conducting a narrative synthesis in a methodical way is outlined by Popay et al. (2006). Initially, this model purports mapping out characteristics (e.g., aims, sample, and assessment measures) and overall findings for each article. Following this, it is considered necessary to explore the details in the data presented (i.e., how data were analysed and interpreted) and consider themes within the findings across studies. The model further highlights the need to scrutinise findings and to note possible contradictions within the data. Lastly, Popay et al. (2006) suggest the need to provide an overview of the limitations and potential biases of the review.

Results

As mentioned above, the final ten articles were quality assessed. A cut of 70% or higher was applied to all articles. All articles scored above the cut-off and, as such, were all included in the review. The data extraction form (Appendix D) was used in order to summarise the characteristics of the ten studies (Table 2) and descriptive information from the ten studies is summarised below. Furthermore, in order to address the second aim of the current review, a summary of the specific DbTs as outlined in each paper is provided in (Table 3). Lastly, a narrative synthesis of the information extracted from the studies is provided.

Table 2

Characteristics of Included Studies

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
Blacker, Watson, & Beech (2008)	*Aim:* To explore the impact of the DbT on anger levels in a prison population. *Design:* Quantitative	N = 62 Male offenders from 6 prisons in the UK Mean age: 31years Offences included: violence (assault with a weapon). Further specific offence related information was not reported in the article.	*Measures:* State-Trait Anger Expression Inventory-2 (STAXI-2; Spielberger, 1999) administered pre- and post DbT. *Analysis:* Results were analysed using t-tests and Multiple Analysis of Variance, and Analysis of Variance (MANOVA and ANOVA) (State Anger, Trait Anger, Anger Expression, Anger Control).	Levels of anger were statistically significantly lower post intervention across all subscales of the STAXI-2. Outcomes showed a significant difference in STAXI-2 scores for State Anger [$F(4,48) = 9.41$, $p < 0.0001$], Trait Anger [$F(3,49) = 12.36$, $p < 0.0001$], Anger Expression [$F(2,50) = 12.48$, $p < 0.0001$], Anger Control [$F(2,50) = 15.89$, $p < 0.0001$], and Anger Index [$F(1,51) = 44.35$, $p < 0.0001$].	• Participants may have been receiving other treatments concurrently • Social desirability factors in the post DbT measures as therapists administered the self-reported STAXI • Reduction in anger and aggressive feelings was shown, however not necessarily translating to a reduction in violent behaviour • Only used self-reported measures, which may have limited a thorough account of the effects of treatment	90%
Colquhoun, Lord & Bacon (2018)	*Aim:* To explore the recovery of British male	N = 5 Aged between 25-50	*Measure:* Semi-structured interviews.	Themes in the data included: not being who they (the participant) used to be; acquiring new perspectives; pro-social	• Small sample (although acceptable for IPA) so not generalisable • Included only offenders with a diagnosed mental	70%

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
	sexual offenders with a mental health (MH) issue who undertook the DbT, and to use this understanding to inform effective treatment delivery and programmes for male sexual offenders with a MH diagnosis *Design:* Qualitative	British male sexual offenders with a mental health (MH) diagnosis (including paranoid schizophrenia).	*Analysis:* Interpretive Phenomenological Analysis (IPA).	relationships; and obstacles to change. The authors posited that obstacles to change were linked to poor memory. Participants understood factors that led to their conviction/hospital admission. Participants gained insight into barriers to change negative behaviour patterns. Participants were able to consider their backgrounds and found insight into how patterns repeat over the years and how behaviours can be learned from a young age. Participants pointed out that practicing pro social communication in hospital can help with future relationships outside of hospital.	health diagnosis. Other groups of offenders (e.g., those with learning difficulties) were not represented in the sample	
Gordon, Shenar, &	*Aim:* To explore the	N = 4	*Measure:* Interviews, narrative accounts.	The emergence of the inner clown helped participants to	• Results not generalisable	70%

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
Pendzik (2018)	process of recovery and rehabilitation with an offending sample. To explore a sphere of drama therapy – the inner clown therapy. *Design:* Qualitative	Ages: 50, 44, 54, the final participant, female –in her late thirties. 3 Males, 1 Female Eastern European, Israeli, Moroccan. Participant No 4 – ethnicity not reported Participants with diagnoses of: Mental Health, Personality Disorder, Intellectual Disabilities. Offences recorded - Theft, addiction.	*Analysis:* Case study	manage conflict, reduce depressive periods, and increase independence. It was reported that the programme can help provide a voice for people struggling to assert themselves (i.e., those who are socially isolated/introverted and who struggle to communicate their emotions effectively).	• Participants could be undergoing other treatment which would confound the results of the study	
Harkins, Pritchard, Haskayne, Watson, & Beech (2011)	*Aim:* To investigate the impact of a combined CBT and DbT intervention in prisons. To explore the	N = 76 55 male, 21 female. Offenders approaching release. Delivered in 11 prisons, 20 occasions	*Measures:* Pre and post intervention STAXI Trait and State Anger using the State-Trait Anger Expression Inventory-2 (STAXI-2; Spielberger, 1999).	Significant improvements were found for all psychometric measures. Self-efficacy, $t(59) = 3.6$, $p < .001$. Motivation to change significantly improved from day one to the last, $t(46) = 2.5$, $p = .015$. Self-reported	• Short length of follow-up • No control group. Not possible to gain access to a control group of participants. This was however an exploratory study intended to provide preliminary evidence of	85%, 80%

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
	impact of this intervention in regard to motivation to change and self- efficacy. *Design:* Mixed-methods	across the UK and Wales. Age range 23-57, Mean age 35 Ethnicity not reported Specific information regarding participants offending histories not reported in the article.	Self-efficacy measured using the General Perceived Self-Efficacy Scale (Jerusalem & Schwarzer, 1992). Motivation to change was measured using the University of Rhode Island Change Assessment (URICA; DiClemente & Hughes, 1990). Confidence in skills was measured using the Skills Rating form from the Geese Theatre Handbook (Baim et al., 2002). Group behaviour and engagement was measured using the Evaluation of Behaviour in the Group form from the Geese Theatre	confidence in skills, $t(35) = 3.6$, $p < .01$. Interviews with participants showed that the programme had a positive impact. Participants reported that perceived self-efficacy, motivation, and confidence in skills following intervention had increased. Further, they reported feeling more able to recover if setbacks were presented.	the impact of the Re-Connect programme • Low internal validity (i.e., the results may be influenced by another factor such as other interventions that participants may be taking part in) • Self-report measures were used which are subject to bias	

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
			handbook (Baim et al., 2002). Semi-structured interviews. *Analysis Quantitative:* Paired *t* tests were used to measure differences between pre- and post-intervention measures. Repeated measure ANOVAs were used to measure change in behaviour and engagement in treatment. *Analysis Qualitative*: no formal qualitative analysis was used.			
Keulen-De Vos, Van Den Broek, Bernstein, Vallentin, & Arntz (2017)	*Aim:* To investigate the impact of a combined Schema Therapy with	N = 9 Males with a diagnosis of personality disorder (PD). A Dutch cohort.	*Measures:* A Multi-Subjects' format was selected using paired (independent comparisons) samples t-tests to measure each therapy component with	Findings showed significantly greater emotions in the family table session from the start of the intervention to when the intervention progressed, eliciting more Vulnerable Child mode (M = 1.88, SE =	• The non-significant results found for the Anger-mode experimental session, could have been due to a low sample size • They used only one outcome measure (MOS),	75%

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
	DbT techniques *Design:* Quantitative (pilot study)	Mean age 38. All patients had an offending history and substance addiction.	Anger Scores and Emotional Vulnerability. The Mode Observation Scale (MOS; Bernstein et al., 2009) looks at 'schema modes' through observation within clinical scenarios in psychiatric/forensic population. *Analysis:* Data were analysed using Analysis of Variance (ANOVA)	0.28) when compared to the baseline score (M = 1.0, SE = 0.006, t(7) = −3.13, p = 0.017). Participants did not show significantly more Angry Child mode after the intervention (M = 1.25, SE = 0.25) commenced compared to pre intervention (M = 1.00, SE = 0.00, t(7) = −1.00, p = 0.35). Participants showed significantly greater Vulnerable Child mode following the intervention commencing (M = 2.06, SE = 0.30) compared to pre intervention (M = 1.09, SE = 0.06, t(7) = 3.26, p = 0.014). Participants showed more Angry Child mode in the	which is an observational measure • The 'anger' session may not have elicited enough provocation to constitute anger. Participants may experience certain emotions in different ways and to different degrees (i.e., that which may be anger provoking for one person may not lead to anger provocation for another) • Different therapists led different sessions, which may have impacted on the outcomes. As there were few sessions, the inconsistency of the potential different approaches by therapists may have influenced the outcomes • A few participants were on anti-depressants during the intervention period which may have	

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
				vulnerability session once the intervention had commenced (M = 1.56, SE = 0.26) than prior to evoking the emotional response (M = 1.00, SE = 0.003, t(7) = 2.18), this finding just missing significance (p = 0.06). Participants did not show significantly more angry-child mode following the intervention commencing, intended to evoke this emotional state (M = 1.17, SE = 0.12) was initiated than pre induction (M = 1.00, SE = 0.008, t(8) = 1.41, p = 0.19). Participants showed significantly more Vulnerable Child mode in the anger session following the intervention commencing (M = 2.33, SE	dampened their emotional expression • The sample included a range of PD categories, the authors suggest using a more homogeneous group for further analysis	

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
				= 0.26) compared to the start of the intervention (M = 1.00, SE = 0.006, t(8) = 5.06, p = 0.001). When comparing emotional vulnerability in all experimental sessions to the remaining sessions, participants showed greater intensity of emotion (p < 0.0125). Participants showed greater intensity of emotional vulnerability in the vulnerability induction session (M = 2.06, SE = 0.31) than in the end session (M = 1.25, SE = 0.13, t(7) = −2.88) however this did not reach significant significance. Participants showed significantly more Vulnerable Child mode in		

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
				the anger-induction session (M = 2.33, SE = 0.26), than in the end session (M = 1.22, SE = 0.12, t(8) = 3.73, p = 0.006). Participants in the family table session did not show more emotional vulnerability compared to the other four sessions.		
Koiv, & Kaudne (2015)	*Aim:* To evaluate the impact of an integrated arts therapy programme on young female juvenile delinquents' emotional and behavioural problems in the criminal	N = 29 Females, Estonian Offenders. Aged 14-17, Mean age 16. Offences noted: drug possession, robbery, violence.	*Measures:* The self-reported version of Strengths and Difficulties Questionnaire (SDQ) and modified Behaviour Checklist (BC) pre and post intervention. The SDQ was used to screen for emotional, conduct and peer problems, hyperactivity and prosocial behaviour. *Analysis:* Independent and paired sample t-tests	T-test analysis revealed significant differences (pre and post test scores) for the intervention group for prosocial behaviours (t(12) = 1.71*, p = 0.05); conduct problems (t(12) = 2.98*, p = 0.05); emotional problems (t(12) = 1.74*, p = 0.05). Significant differences for these three constructs following intervention (between the intervention and control group) were further found: prosocial	• Results cannot be generalised to other offender populations • Sample size was small for a quantitative study • It was not possible to randomly assign participants to groups – making it problematic to rule out bias	85%, 80%

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
	justice system *Design:* Quantitative (quasi-experimental)			behaviours (t(12, 17) = 1.71*, p = 0.05); conduct problems (t(12, 17) = 3.99**, p = 0.01); emotional problems (t(12, 17) = 1.75*, p = 0.05). For constructs of hyperactivity and peer problems, no significant differences were found in pre to post scores for both groups (t(12) = 1.30, p = < .05); t(17) = 0.24, p = < .05). Significant changes were noted in outcomes measured by the SDQ between the intervention and control group: reduced conduct and emotional problems (t(12, 17) = 3.99, p = 0.01; t(12, 17) = 1.75, p = 0.05), and improved prosocial behaviour (t(12, 17) = 1.72,		

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
				p = 0.05) in the intervention group. For the BC measure, significant differences were found in aggressive (t(12) = 4.79, p = 0.01), and prosocial behaviour (t(12) = 1.93, p = 0.05), but not in withdrawn behaviour between pre- and post-intervention scores (t(12) = 0.99, p = >.05), as well as between the intervention group and control group scores following the intervention. Decreases in incidents of aggressive behaviour was found in the intervention group as was increased prosocial behaviours for the intervention group (prosocial – t(12, 17) = 2.49, p = 0.01); aggressive – t(12, 17) = 6.28, p = 0.05;		

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
				withdrawn - (t(12, 17) = 0.14, p = <.05).		
Mundt Mundt, Marín, Gabrysch, Sepúlveda, Roumeau & Heritage (2019)	*Aim:* To explore the effects of being involved in a theatre project, for people with MH issues and offending histories. *Design:* Qualitative	N = 5 Three prisoners with diagnoses of depression and associated substance misuse issues, and at risk of suicide. Two were the parents of deceased prisoners (whom had an offending history) 4 males, 1 female 3 males had committed violent crimes Aged between 36 and 65, mean age 46.	*Measures:* The Mini-International Neuropsychiatric Interview (MINI) was used to assess mental disorder and drug use. Semi-structured interviews *Analysis:* Thematic analysis	Participants reported that they felt socially criticised when performing theatre. However, they reported a freedom in challenging social and societal norms and felt liberated in rising above repression. Participants reported that the intervention helped them communicate more effectively. There was a clear sense of achievement from the participants. The public responded positively to the performance which led to prisoners being able to envision a new self, and to feel motivated for a different future.	• Small sample size • Non-homogenous participants; differing criminal justice involvement, as well as differing backgrounds in terms of mental health experiences • Unable to conduct a second interview with some of the participants due to a number of factors out of the researchers' control	85%
Reiss, Quayle, Brett &	*Aim:* To investigate the impact of	N = 12	*Measures:* Pre and post trait and state anger using the State-Trait Anger	Significant improvements on both the 'how angry' and the 'how react' (z = 2.31, p	• No control group • Lack of ethnic diversity within the sample	90%

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
Meux (1998)	a week-long theatre project on anger levels. *Design:* Quantitative	British male patients/young offenders with a MH/PD diagnosis, in a secure hospital with violent histories Mean age: 26 Offences noted: homicide, sexual offences, and violent behaviours.	Expression Inventory-2 (STAXI-2; Spielberger, 1999). A five-point scale, a maximum score of four equalling greater benefits. *Analysis:* Wilcoxon matched-pairs signed-ranks test to compare differences pre/post theatre project.	= 0.02; z = 2.09, p = 0.04) scale following completion of the programme. No significant differences between scores immediately following the programme and after a 3-month follow up period. However, prior to follow-up differences were statistically significant ('how angry' z = 2.75, p = 0.006; 'how react' z = 2.39, p = 0.02). No statistically significant changes were found on the State-anger scale. The Trait-anger scale however showed a significant improvement from pre to follow-up (z = 2.24, p = 0.02) but not from pre to post, and post to follow up. There were trends in the data showing improvement however this	• More follow up measures would be necessary if the study were to be replicated • Social desirability in the feedback, as self-reported measures were used • The findings do not necessarily translate to a reduction in violent reoffending	

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
				failed to reach statistical significance. The AX/con scale however, showed a significant increase from pre to follow-up ($z = 2.65$, $p = 0.01$) but the initial benefits/changes, in spite of higher score, were not significant individually.		
Stahler (2007)	*Aim:* To describe and discuss the experiences of participants of a DbT programme Design: Qualitative	N = not specifically recorded, a small sample taken from Groups of 8-12 participants. Female prisoners in recovery/offenders in recovery (addiction) Ethnicity not specified, one described as bi-racial	Measure: Semi-structured interviews Analysis: Narrative accounts of participant experiences were analysed and summarised	Participants reported that the intervention helped them explore their identities outside of their addictions, and to come to terms with static elements of themselves and how they can think differently about these elements. It helped them deal with issues around stigmatisation (addict and prisoner). They reported feeling more optimistic about their futures.	• The authors omitted to provide any reflections/ limitations of the study. • No structured qualitative methodology was used. • A lack of details regarding the participants.	95%, 80%

Authors	Aim and design	Sample and setting	Measures and analysis	Findings	Limitations	Quality score
Stephenson & Watson (2018)	*Aim:* To investigate whether the DbT programme increased levels of hope and mental well-being. In addition, to explore the views of participants regarding the programme. *Design:* Mixed-methods	N = 21 Female offenders – some with MH diagnoses and/or substance misuse histories Aged 20 - 49, mean age of 31 Specific information regarding offending behaviour was not provided, however, some of the sample were believed to have suicidal ideations. Mixed ethnicity: African, Asian, Indian, Caribbean, Asian, Pakistani, British, Irish	*Measures:* Pre and post intervention measure using the Beck Hopelessness Scale (BHS) and the Warwick-Edinburgh Well- being Scale (WEMWBS)	Significant reduction in scores was evident from the BHS pre- post time point, with related high levels of hope and reduced suicidal ideation(t (20) = 5.58, p = < .001, d = 1.26). The WEMWBS showed significantly higher scores which relates to a positive effect on mental well-being, problem solving, and being able to communicate with others more effectively (t (20) = -4.50, p = < .001, d = 1.18).	• Relatively small sample for the quantitative element of the study. • No follow-up to see if the effects were long lasting • A non-homogenous sample due to difficulties in recruitment. • Prisoners may have been participating in other programmes • Self-report measures may be subject to bias	95%, 90%

Descriptive Overview of Results

Four quantitative papers (Blacker et al., 2008; Keulen-De Vos et al., 2017; Koiv & Kaudne, 2015; Reiss et al., 1998), four qualitative papers (Colquhoun et al., 2018; Gordon et al., 2018; Mundt et al., 2019; Stahler, 2007), and two mixed-method papers (Harkins et al., 2011; Stephenson & Watson, 2018) were included in the review.

Samples

The majority of articles specified how many participants their research was based on. In total, the number of reported participants involved in this review was 223 (across all papers). However, one paper (Stahler, 2007) did not specify how many participants took part. The majority of participants across studies were male (N=150). However, three studies described mixed gender samples (Gordon et al., 2018, N = 4, male *n* = 3, female *n* = 1; Harkins et al., 2011, N = 76, male *n* = 55, female *n* = 21; Mundt et al., 2019, N = 5, male *n* = 4, female *n* = 1) and two studies used an all-female sample (Stephenson & Watson, 2018, N = 21; Koiv & Kaudne, 2015, N = 29).

The studies took place in a variety of locations and included participants from the UK (Blacker et al., 2008; Colquhoun et al., 2018; Reiss, et al., 1998; Harkins et al., 2011; Stahler, 2007; Stephenson & Watson, 2018), Holland (Keulen-De Vos et al., 2017), Estonia (Koiv & Kaudne, 2015), Israel (Gordon et al., 2018), and Chile (Mundt et al., 2019).

Psychiatric and Offending History

All of the studies took place in a prison or secure environment, with the exception of two which took place in a community setting (Keulen-De Vos et al., 2017; Stahler, 2007).The samples from the studies included in this review ranged in their offending histories and mental well-being (Table 2). Blacker et al. (1998) targeted males with a violent conviction and anger issues (N=62). Colquhoun et al. (2018) included participants from a secure hospital, all male and over 18. The participants in the Gordon et al. (2018) study had historic

substance misuse issues and a diagnosis of personality disorder. The study by Harkins et al. (2011) included prisoners who were close to release from prison. Keulin et al.'s (2017) study included male service-users within a psychiatric facility; all of whom had a diagnosis of cluster B Personality Disorder[1]. Koiv and Kaudne (2015) reported that all participants were young offenders in a high security correctional institution. This institution was a long-term residential placement. Similarly, Reiss et al. (1998) included participants from a juvenile ward from Broadmoor hospital (a high security hospital). Stahler's (2007) study targeted female prisoners in recovery from addiction. Stephenson and Watson (2018) reported that participants were females in a prison setting.

Type of DbT used

All studies provided a brief description of the type of Drama based Therapy (DbT) intervention which was being evaluated in their research. Two studies used an intervention which they described as adopting a Drama based therapeutic approach (Harkins et al., 2011; Mundt et al., 2019). One study described their intervention as a Drama based programme (Blacker et al., 2008). Six studies used the term Drama Therapy (Colquhoun et al., 2018; Gordon et al., 2018; Keulen-De Vos et al., 2017; Koiv & Kaudne, 2015; Reiss, et al., 1998; Stahler, 2007). Lastly, one study was described as an applied theatre intervention (Stephenson & Watson, 2018). (Table 3) for a more detailed description of each type of DbT used.

[1] Cluster B personality disorders include antisocial personality disorder, borderline personality disorder, histrionic personality disorder and narcissistic personality disorder (Keulin et al., 2017).

Table 3

Types of DbT

Paper	Study	Type of Drama based Therapy (DbT)
1	Blacker, Watson, & Beech (2008)	**Term used**: Drama based programme **Description of programme:** The programme (*Insult to injury*) was developed by the Geese Theatre Company with the aim of enabling "offenders to identify and generate strategies and skills for dealing with potentially volatile situations, and to provide a safe and supportive environment in which to practice and evaluate these strategies" (p 129). The drama based approach was combined with cognitive behavioural anger management principles. Regarding the purpose of the programme, the author's noted that both dramatic and theatrical techniques were used to "help inmates with problems of emotionally-driven violence to identify strategies and skills for dealing with potentially volatile situations in an appropriate way" (p.132). The programme was split into three blocks. The first exploring internal processes which impact upon behaviour. The second explored "…masculinity, power and control, pride and shame, the consequences of anger related behaviour for others and victim awareness" (p.132). The third aimed to consolidate learning and provided participants with the opportunity to consider choices they make and alternatives to their behaviour. In addition, they were given the opportunity to discuss more effective/positive ways of reacting to situations and to rehearse those skills through role play. There was an onus on reflection throughout.
2	Colquhoun, Lord & Bacon (2018)	**Term used**: Dramatherapy **Description of programme:** The Dramatherapy component was facilitated by a qualified therapist and used techniques such as role-play, enacting narratives, guided reflection, sculpting perspectives, observation, and commenting on others. It was delivered as a 3-day Dramatherapy workshop, which was part of an overall 18-week treatment for sexual offenders with a mental health diagnosis (MH) (including paranoid schizophrenia). The

		intervention itself used the Good Lives Model approach (Andrews, Bonta, & Wormith, 2011). The aims were to address empathy and emotional awareness. It aimed to help offenders use their lived experiences and apply that to their recovery.
3	Gordon, Shenar, & Pendzik (2018)	**Term used:** Dramatherapy approach (Clown therapy) **Description of programme:** A Drama-Therapy approach to working with individuals with addiction issues was developed whereby clients were led to discover their *inner clown*. Using drama therapy techniques, each client's key/dominant and oppositional personality roles were recognised and brought together through embodiment, leading to the emergence of the clown. Clown therapy is referred to in the paper as a branch of dramatherapy which focuses on how one can activate *the inner clown* and can behave extrovertly. Participants are required to: actively participate (role-play); be part of an audience; and support one another. The therapist is also required to participate in the clown role. The authors describe three levels of therapy. The first, to identify with an animal and to mimic their behaviours and movements; letting go of inhibitions, using their bodies to explore their animal character. The second stage moves to developing *the clown*. Participants are encouraged to characterise their personalised clown. Stage 3 purports that the clown has been shaped and should be explored in different situations (i.e., interacting and communicating with others, exploring emotions and acknowledging what is needed to feel equipped in dealing with emotions through this new role/persona).
4	Harkins, Pritchard, Haskayne, Watson, & Beech (2011)	**Term used**: Drama based intervention **Description of programme:** A three-day intensive programme using CBT and drama techniques. The intervention was developed by the Geese Theatre Company (GTC). Performance, role-play, and mask work were used as techniques. Within the programme, participants explored issues/situations related to inmates' release, preparing them for life outside prison. There was a focus on active group work to encourage communication and putting trust in others. GTC facilitators presented scenarios and helped inmates engage with characters in the performances. Still images were used as prompts to create potential high risk scenarios. GTC actor practitioners helped the participants manage situations through the use of role-plays. Responses to challenging situations were played out and reactions using adaptive behaviours were role-played. The programme aimed to help participants to learn and develop tools to manage their behaviour. Participants acted out different situations and played an opposite role to their default role, which could be helpful when faced with challenges in the community.

5	Keulen-De Vos, Van Den Broek, Bernstein, Vallentin, & Arntz (2017)	**Term used**: Drama therapy techniques **Description of programme:** The Drama-Therapy techniques were used in eliciting emotional vulnerability in PD clients. The drama therapists used realistic play techniques/exercises to evoke anger and vulnerability emotions. The programme consisted of five sessions each with their own protocol: an introduction session, a general experiential session, a session to evoke emotional vulnerability, a session to evoke anger, and a wrap-up session. Part 1 (session 1) involved an introduction, the stating of rules, and play versus reality principles were explained. Basic emotions were outlined (i.e., anger, happiness, sadness, and anxiety). Participants were then asked to pick cards which show an emotion. The participant and the therapist then played out this emotion. Part 2 involves the *family table* where the participant describes their childhood home situation with family members/caregivers that were around at this time. The participant is encouraged to visualise their family around a table and are invited to talk about how they behaved towards the participant at that time. A home setting is created with props; used to mirror memories with the view of initiating underlying emotions. Part 3 involves evoking emotional vulnerability; the therapist now has a picture of what needs have not been met in the past for the participant. The participant is asked for an example of such a situation where their need was not met. The therapist then plays the person that did not meet the participant's need. Following this, the participant is asked about their emotions at this time. They then switch roles and re-enact the scenario, however a new situation is created where the participant's need is met. The objective being to expose underlying emotional vulnerability. Reflection follows each part, and opportunities for breaks. Part 4 involved evoking anger; the participant is asked about a scenario that elicited anger. The therapist and participant both talk about the situation, who was there and what happened. The participant then acts out the *angry* part, and the therapist is in the role of the object of anger. The scene follows a realistic protocol in as much as the participant is to embellish what really happened. Then the two switch roles with the objective of the participant being to release blocked anger. Part 5 involves *wrap up* with both participant and therapist reflecting on earlier sessions, and the participant is encouraged to express how they experienced this.
6	Koiv & Kaudne (2015)	**Term used**: Drama therapy intervention **Description of programme:** The drama therapy component was part of a wider arts based therapy intervention whereby a range of creative therapies were used (i.e., music and dance). The drama therapy component made up 30% of the overall intervention. Closely linked was the movement therapy element of the intervention which focused on the use of body movement to find connections between mind and body. Pre-structured and improvised drama, storytelling, and role-play drama therapy aiming to encourage the expression of emotions and the

		development of positive self-concept was employed. More specifically, the aim was to help with expressing emotions in a prosocial manner, to learn emotional regulation skills, and to develop adaptive communication skills.
7	Reiss, Quayle, Brett & Meux (1998)	**Term used**: Dramatherapy project **Description of programme:** The dramatherapy was described as group based, interpersonal, interactive and educational. It works through intense, emotional and dynamic confrontations which, because they reflect or re-enact the subjects' own experiences, have considerable interpersonal reality. The programme was a week-long therapeutic theatre project aimed at reducing levels of anger. The project was a Geese Theatre Company (GTC) drama based intervention with CBT techniques programme, called *The Violent Illusion Trilogy*. Through conducting plays and workshops, Actor Practitioners act out violent scenarios. Participants were encouraged to explore cognitive processes that underly their behaviour. The GTC team act out destructive behavioural cycles and teach anger management and problem-solving skills, promoting prosocial behaviours for the future. The program aimed to teach participants how to replace anger with more effective ways of communicating distress.
8	Mundt, Marín, Gabrysch, Sepúlveda, Roumeau, & Heritage (2019)	**Term used**: Drama based therapeutic approach **Description of programme:** The programme was described as a theatre project for mentally disordered prisoners. The project involved participants performing a Shakespearian piece. The participants were three prisoners who had a diagnosis of depression, and associated substance misuse issues. The programme described Drama-Therapy workshops in line with rehabilitation; using drama and movement production as a means to communicate with others and the wider community. The paper described the programme as using dramatic expression to channel the unconscious mind. Participants used their own stories and experiences in performing and understanding their own behaviours. The roles within this Shakespearean piece involved trauma and violence, to which prisoners can relate. The use of Shakespeare's language gave participants new and different ways to communicate.
9	Stahler (2007)	**Term used**: 'Prayerformance': A Drama Therapy Approach **Description of programme:** The programme was delivered over a 12-week period. The programme used a spiritual and psychological model of addiction. Dramatherapy techniques such as role-play, storytelling, character development, and making and using masks were used. The paper described addiction as a survival and escape

		mechanism, and state that regardless of the nature of addiction (e.g., substances, gambling) underlying internal conflicts (i.e., the inner self; one's connection with others; and the spiritual self) may be relevant. The programme uses a rehabilitation model aimed to help with the release of tension in the body through movement; encouraging participants to physically express their inner world. This intervention described a mirroring process, which involves exploring oneself outside of their addiction. The participants are coupled up and sit knee to knee. The dyads discuss how they are feeling and communicate their bodily experiences. Therapists initially work on trust within the group which facilitates the therapeutic partnership of the dyads. Participants are encouraged to form relationships, and to focus on the process of the spiritual connection in their groups. Each participant takes turns to lead the group. Role-play exercises aim to promote health and healing by exploring ways to manage and live with parts of oneself they may dislike. In addition, masks are used to represent the participants' hopes and dreams for the future.
10	Stephenson & Watson (2018)	**Term used**: Applied theatre intervention **Description of programme:** The intervention (named *Scratching the Surface*), developed by Geese Theatre Company (GTC), aimed to enable offenders to learn and practice new skills and consider these within their social context. The intervention further aims to help participants gain insight into their thinking patterns and how these impact upon their behaviour, and to consider the roles that they play in their lives whilst providing them with assurance that they can create new, prosocial roles. The programme used the metaphor of a *mask*, with the mask symbolising the façade that others see, where real thoughts and feelings are hidden. Short performances were conducted by GTC's Actor Practitioners and participants are encouraged to engage in tasks such as role-plays. Through acting out difficult situations, programme facilitators help participants learn how to deal with anger, and teach skills in communicating, problem solving, and coping. A key aim was to increase confidence. Part of the work of the GTC is tailoring concepts to individuals, encouraging self-reflection and providing memorable images/moments. The programme concluded with participants doing their own performance, which had been tailored to their specific needs by facilitators.

Psychometric Measures

In terms of the measures used in the quantitative studies, Blacker et al. (2008) used the State-Trait Anger Expression Inventory-2 (STAXI-2; Spielberger, 1999) pre- and post DbT. This was administered and collected by Geese Theatre staff. The STAXI-2 was also used in studies by Reiss et al. (1998) and Harkins et al. (2011). Harkins et al. (2011) further used the General Perceived Self-Efficacy Scale (Jerusalem & Schwarzer, 1992). The GSE is used to assess hypothetical self-beliefs in dealing with challenges in one's life. The University of Rhode Island Change Assessment [2] (URICA; Diclemente & Hughes, 1990), the Skills Rating form in the Geese Theatre Handbook (Baim et al., 2002), and the Evaluation of Behaviour in the Group form from the Geese Theatre handbook (Baim et al., 2002). Keulen-De Vos et al., (2017) used The Mode Observation Scale (MOS; Bernstein, de Vos, & Van den Broek, 2009). Such measures have been validated for use within these settings.

Koiv and Kaudne (2015) asked participants to complete the self-report versions of the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) and the Modified Behaviour Checklist (BC; Ross, Lacey and Parton, 1965) before and after attending the intervention. Stephenson and Watson (2018) used the Beck Hopelessness Scale (BHS) and the Warwick-Edinburgh Well-being Scale (WEMWBS; Stewart-Brown & Janmohamed, 2008) to measure change following intervention.

In terms of the measures used in the qualitative and mixed design studies, Stahler (2007) used narrative accounts of participant experiences. Stephenson and Watson (2018) used semi-structured interviews, as did Colquhoun et al. (2018) and Harkins et al. (2011). Mundt et al. (2019) used the Mini-International Neuropsychiatric Interview (MINI; Sheehan et al., 1998). The researchers approached participants and conducted diagnostic and

[2] Outcomes on the URICA inform changes to motivation, which can help to inform treatment pathways (Diclemente & Hughes, 1990).

qualitative interviews. Gordon et al. (2018) used interviews, narrative accounts, and case studies to explore the views of participants.

Analysis of Findings

In terms of the analysis used in the quantitative studies, Blacker et al. (2008) analysed their data using *t*-tests, Multiple Analysis of Variance (MANOVA), and Analysis of Variance (ANOVA) to determine whether anger levels reduced following the intervention. Harkins et al. (2011) also performed *t*-tests and repeated measure ANOVAs in order to measure differences between scores collected pre- and post-intervention. Keulen-De Vos et al. (2017) used a multi-subjects format using paired (independent comparisons) samples *t*-tests and ANOVAs to measure each therapy component with Anger Scores and Emotional Vulnerability. Koiv and Kaudne (2015) also used *t*-tests (Quasi-experimental design, with controlled samples). Reiss et al. (1998) used a repeated measurements design (Wilcoxon matched-pairs signed-ranks test, to compare differences over time). Stephenson and Watson (2018) also performed repeated measures *t*-tests to look for significant differences between scores obtained pre- and post-intervention.

In terms of the analysis used in the qualitative and mixed-methods studies, Colquhoun et al. (2018) used Interpretative Phenomenological Analysis (IPA; Smith et al., 2009), and Stephenson and Watson (2018) and Mundt et al. (2019) used Thematic Analysis (Braun & Clarke, 2006) to analyse transcripts. It is of note that Stahler (2007), Harkins et al. (2011), and Gordon et al. (2018) omitted to specify the type of qualitative analysis used in their studies (this is reflected in the quality assessment scores).

Recruitment of Participants

Regarding participant recruitment for the included studies, most authors omitted to describe who approached potential participants (Keulen-De Vos et al., 2017; Harkins et al., 2011; Reiss et al., 1998; Stahler, 2007). In the Colquhoun et al. (2018) study, it was reported

that the first author targeted participants from a Sex Offenders Group. Gordon et al. (2018) described a compulsory intervention (to avoid prison), and some attended voluntarily, but omitted to describe how potential participants were approached. In the Koiv and Kaudne's (2015) study, it was stated that all inmates who were attending the DbT voluntarily were invited to participate in the study.

Stephenson and Watson (2018) reported that participants were recruited through a safer custody team within prisons. Mundt et al. (2019) reported that one of the authors recruited participants. Blacker et al. (1998) reported that participants were approached by prison staff in order to ascertain whether they would be willing to take part.

Narrative synthesis of findings

As noted above, there was considerable variation between studies in terms of the type of DbT being evaluated, the study design, the measures used, and the analysis conducted. However, it was felt that there were sufficient commonalities between the DbTs used and the focus of the individual studies to be able to collate themes from the findings of the studies.

Themes derived from study outcomes

Anger and Aggression. Six DbT interventions placed a focus on reducing anger or aggression, and as such, measured or explored the impact that the intervention had on anger or aggression of the participants. Studies 1 (Blacker et al., 2008), 4 (Harkins et al., 2011), 5 (Keulen-De Vos et al., 2017), 6 (Koiv & Kaudne, 2015), 8 (Reiss et al., 1998) and 10 (Stephenson & Watson, 2018) all described a reduction in anger and aggression as a result of their respective DbT.

More specifically, Study 1 described a significant reduction in levels of anger following the implementation of the DbT programme, and also reported that incidents of anger related behaviour had reduced. STAXI outcomes showed significant improvements for all scales. Further, the authors reported an increase in participants using more prosocial anger

management skills post intervention. Participants acquired an understanding of the sequence of how thoughts and feelings impact on behaviour. They learned about making choices and adaptive problem-solving skills as an alternative to expressing anger. Similarly, Study 10 reported that some participants felt that taking part in the DbT had led to them thinking before they acted.

In Study 4, the authors concluded that the DbT may have resulted in reductions in levels of anger expression (as measured by the STAXI). In addition, data taken from interviews conducted post intervention with a sub-set of participants to ascertain their thoughts on the programme revealed that some participants felt that they would be able to manage anger/aggression in the community more effectively with the self-manage tools they had learned from the DbT.

Study 5 looked at anger scores and emotional vulnerability following each intervention component. Differences in scores before compared to post programme were not significant, suggesting that the DbT programme did not have an impact on the participants. However, the authors noted that the power of the initial data analysis was weak given that it was a pilot study with a small sample. They further noted the breadth of the term emotions, and that the 'anger' session may not have elicited enough provocation to constitute anger. Indicating that people experience anger and aggression in different ways, provocation to anger may vary across different forensic populations. Therefore, outcomes should be treated with caution regarding improvement reported for levels of anger and aggression.

The findings of Study 6 were concurrent with those of Blacker et al. (2008) and Harkins et al. (2011). The intervention was found to be effective in reducing emotional and behavioural problems such as aggression, and in strengthening pro-social behaviours. A comparable control group was used in order to provide meaning to the outcomes of the experimental group. Post-intervention measures looked at levels of incidents of aggression

and other behavioural issues. Findings showed significant reductions in conduct and emotional problems and improved prosocial behaviour for both the experimental and control groups. The control received no interventions, they merely continued with their normal daily activities. The authors concluded that the integrated Drama-Therapy intervention was beneficial in reducing emotional and behavioural problems such as aggression, and in strengthening pro-social behaviours in young female offenders.

Furthermore, Study 8 reported that the intervention helped participants learn how to replace anger with more effective ways of communicating distress. They reported participants acquired greater understanding of recognising when anger starts and being more mindful in that moment to be able to take control of their actions. When measuring outcomes in general, it was reported that there was a significant improvement on the trait anger scales. However, closer analysis of the data showed a significant reduction in trait anger but not state anger, and this was maintained at three months follow-up. The authors purport that this may have been due to participants not always reacting with a behavioural response (i.e., young male offenders with a MH/PD diagnosis may be more contained by relational security of the institution, suggesting that this client group may have more chronic anger issues). Therefore, this subgroup may not have taken as much from the intervention as other categories of offenders. Participants with high anger scores may often experience heightened anger feelings, however they may suppress them as opposed to lashing out physically/verbally. There was an increase in the number of attempts to control anger outbursts, as rated by participants on a five-point scale. However, the authors acknowledged that just because a reduction in anger and aggressive feelings was shown this does not necessarily translate to a reduction in violent behaviour.

Hope. Although hope is associated with mental well-being (see the below theme), it was felt necessary to include this topic as a separate theme as increasing levels of hope was the focus of nearly half of the DbT programmes used and, subsequently, was an issue explored by researchers. Studies 7 (Mundt et al., 2019), 9 (Staher, 2007), 10 (Stephenson & Watson, 2018) and 3 (Gordon et al., 2018), describe participants feeling more hopeful in relation to their future/recovery as a result of DbT. Participants in Study 7 reported a freedom in challenging social and societal norms and felt liberated in rising above repression. Participants reported that by the public responding positively to the performance this led to them being able to envision a new self, and to feel motivated for a different future and to feel motivated to change.

Study 9 reported on how the DbT intervention helped participants explore their identities outside of their addictions, and helped them come to terms with static elements of themselves and help them see how they can think differently about these elements. In addition, DbT reportedly helped them deal with issues around stigmatisation (i.e., potentially being treated differently due to being an addict and prisoner). They reported feeling more optimistic about their futures. One participant commented that their mask was a warrior that wanted to succeed in life. Others commented on feeling a sense of freedom and control following the intervention; being more optimistic about their futures. The authors concluded that their DbT intervention was successful with this group of female addicts (N = not recorded).

Study 10 used the Beck Hopelessness scale (BHS; Beck et al., 1974) and reported significant improvements in scores pre- and post DbT. Female prisoners who took part in the DbT intervention were reported to be more hopeful following completion of the programme. Following the DbT, participants were more likely to agree with statements such as “I look forward to the future with hope and enthusiasm”, “when I look ahead to the future, I expect

that I will be happier than I am now", and "I can look forward to more good times than bad times". They were less likely to agree with statements such as "the future seems vague and uncertain to me" and "I just can't get the breaks, and there is no reason I will in the future". Furthermore, in the questionnaire element of the study, some participants reported that, following the DbT, they were better able to think about life goals. One participant commented, "I can see to the future, it looks a lot brighter" (Stephenson & Watson, 2018). Lastly, participants in Study 3 reported that their DbT (clown therapy) had helped them be more optimistic about the future.

Mental Health and Well-being. Studies 3 (Gordon et al., 2018) and 10 (Stephenson & Watson, 2018) describe an increase in mental well-being and a reduction in depressive symptoms in their samples. Study 3 reported that their branch of DbT (i.e., clown therapy) led to a reduction in depression. Participants stated that this was a new experience that had helped them gain insight into the depressive state they had fallen into. For one individual, it was further reported that the intervention had helped them make links between the loss of a family member and them self-medicating with substances to ease pain. The inner clown exercises were reported to help with balancing their inner feelings with their outer image and that it had further helped with depressive periods in their lives. It was noted that clown therapy is not aimed at curing addiction or at specifically preventing relapse, but that it teaches people to express emotions more freely - finding alternative ways to cope with negative emotions and stress. They purported that clown therapy could be effective with a range of psychological disorders and it can provide a voice for people struggling to assert themselves.

Whilst acknowledging the limitations of their research, Study 10 reported that the DbT was effective for participants in terms of improving positive mental well-being as measured by The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS, Tennant et al.,

2007). The WEMWBS measure assesses happiness and life satisfaction states and shares features with the Short Depression-Happiness Scale (SDHS, Linley et al., 2004). As such, improvements in scores could be seen as being associated with lower levels of depression and increase in happiness. High scores on the WEMWBS indicate more positivity in participants' thinking, and their subsequent behaviour. The authors reported that scores recorded post DbT completion revealed participants felt more confident, had higher efficacy, self-esteem and acceptance. On average, participants reported more occurrences of positive feelings, for example, scores increased on items such as "I've been interested in new things", "I've been dealing with problems well", and "I've been feeling good about myself" (Tennant et al., 2007). Outcomes following the programme revealed participants felt more confident, better able to express their feelings to others, and more hopeful for a better future (see previous theme).

Furthermore, as mentioned above, Study 10 reported that results indicated that the DbT was affective for participants in terms of increasing levels of hope. Hope has been found to be negatively correlated with depression (i.e., as hope increases, depression decreases) (Leite et al., 2019), as such, higher levels of hope reported by participants following completion of the DbT, could indicate lower levels of depression in participants. However, the use of a specific depression measure would be needed to confirm this. Further qualitative analysis in Study 10 revealed that participants reported positive responses on Likert scales (i.e., increased confidence, better at opening up, feeling better within themselves), giving further indication of improved well-being.

Communication and Relationships. Studies 2 (Colquhoun et al., 2018), 3 (Gordon, Shenar & Pendzik, 2018), 7 (Mundt et al., 2019), and 10 (Stephenson & Watson, 2018) all reported that their respective DbT intervention led to improvements in the area of communication skills and/or relationships. Study 2 reported themes within their data, one of

which was pro-social relationships whereby participants reported that the DbT had helped them in this area. Participants practiced effective communication skills that they learnt during the DbT in the hospital and believed that this would help with maintaining healthy future relationships outside of hospital. Furthermore, the DbT placed a focus on group dynamics and the barriers within the group, with a few participants initially reporting that they did not like some of the other group members. Reasons given for this was that they were unsure sure how others in the group perceived them, and when asked to contribute within the group, some reported feeling self-conscious. The authors highlighted that a common trend in the data was that participants largely compared themselves to their peers, which may have impacted on their engagement in the workshop. Overall, however, it was found that DbT helped acquire pro-social communication skills in order to build their relationship with other group members and relationships with others outside the group.

Study 3 reported that clown therapy helped participants manage conflict in relationships. As previously stated in the theme relating to mental well-being, it was felt that clown therapy can provide a voice for people struggling to assert themselves and can help people to communicate their emotions to others. Similarly, Study 7 reported participants felt the DbT resulted in them being able to communicate more effectively. Participants reported that - through their physical expression – they were able to highlight social awareness within prisons, and to communicate in a meaningful way with a wider range of people. Lastly, participants in Study 10 reported that they felt the DbT had resulted in them being better able to open up to people. Furthermore, one participant felt that the DbT was particularly useful for improving relationships in a prison setting and, as such, should be delivered more widely to prisoners.

Gaining Insight into Oneself. Participants in Study 2 (Colquhoun et al., 2018) reported that, through the specific DbT used, they had learnt something about themselves. The authors reported that the role-play element of the intervention was effective in that it helped participants understand how avoiding engaging in crime can help them with relationships and to achieve their goals. The authors reported that it is through reflection that offenders connect with the shared experiences of the group. Being able to act out scenarios and watch these scenarios being acted enables the offender to gain insight into their actions and thought processes.

Similarly, in Study 3 (Gordon et al., 2018), one participant stated that DbT helped them gain insight into their depressive state; helping them see the impact of a bereavement and subsequent self-medicating leading to substance misuse.

A Sense of Achievement. Study 10 (Stephenson & Watson, 2018) reported that the DbT intervention led to participants reporting that they felt they had done something they were proud of. Similarly, Mundt et al. (2019) (Study 7) reported that there was a clear sense of achievement noted from participants. The authors reported that the sense of achievement had led to an improvement in participants' self-image, and participants "expressed the experience of personal success" (p. 8), and that they believed this to be linked to the wider concept of social inclusion. Some participants reported that they had initially felt like being on display and open to being socially critiqued when performing. They reported a freedom in challenging social and societal norms and felt liberated when performing. The authors reported the joy that was observed in putting on a performance, and the overt positive feelings on show from participants. The authors concluded that it is possible to channel the unconscious mind of people with mental health issues through dramatic expression.

Discussion

The aim of this review was to explore the types of DbTs used with forensic populations, and to synthesise the literature regarding the impact of DbT interventions with forensic populations. Methodological limitations notwithstanding (see below for details), the findings regarding the positive impact of DbTs are promising. This is consistent with findings from DbTs used in clinical settings with patients with intellectual disabilities and mental health diagnoses (Feniger-Schaal & Orkibi, 2020). More specifically, DbTs outlined in this review were found to have a positive impact on levels of aggression/anger, hope, mental well-being, and communication. Furthermore, it was reported in some DbTs that participants gained insight into themselves and felt a sense of achievement following programme completion.

As can be seen in Table 3, this review has found there to be a wide range of DbTs used with forensic populations; each having a different focus and approach whilst bearing some commonalities. Authors of the included studies highlighted that key to DbT was using the expression of theatre to model adaptive behaviours. These expressions generally include both the participants and the practitioners being involved in performance.

Each paper outlined what is unique to DbT in general as compared to other programmes delivered to forensic populations. As mentioned, historically and currently, standardised/manualised programmes are used routinely in forensic settings (e.g., secure hospitals, community rehabilitation facilities, prisons). The majority of the authors cited in this review highlight that the key difference between standardised interventions and DbT interventions is the creativity involved with the programme and the flexibility to tailor the programme to the individual population sample. Authors comment on how a particular form of DbT is used to address the needs of a particular group of individuals.

Two core techniques mentioned in the majority of papers are role-play and reflection; a common feature noted across drama based programmes (Gersie, 1996). As mentioned above, due to the evidence base around the use of role-play, this method has also been used in manualised offender behaviour programmes (Antonowicz & Ross, 1994). When outlining the dimensions of effective correctional counselling, Andrews and Bonta (2017) note it to be good practice to use role-play as part of skills building in relation to problem solving and self-management. Where role-play is used in the DbTs reported in this review, it not only serves for the purpose of skills practice but also to provide participants with insight into their own thoughts and behaviours. Whilst manualised offender behaviour programmes may use role-play as a small element of a programme, role-play is a core component in DbT, as would be expected.

Furthermore, providing participants with an opportunity for reflection was key to many DbTs; helping offenders gain insight into their behaviour and be more engaged in their treatment (Frost & Connolly, 2004). Other arts-based interventions with prisoners have also highlighted the importance of reflection. Mullen (1999) emphasised the importance of reflection in the process of behavioural change stating that reflection facilitates learning and development. More specifically, some qualitative feedback from a participant of the Geese Theatre Company's DbT (Baim et al., 2002) stated that the mask exercises were useful in reflecting on themselves and that through the exercises they learnt how they react in different situations and that their decisions can contribute to the outcome of a situation.

The onus on facilitators helping participants reflect on and gain insight into their thoughts and behaviours was mentioned in many of the DbT descriptions (see Table 3). Broadly speaking, findings from this review suggest that DbTs are successful in assisting participants gain insight into their own behaviours. This finding is concurrent with that of other arts based therapies used with forensic populations (Koch et al., 2014; Meekums &

Daniel, 2011). Gaining insight into their behaviours may be a necessary step to recognising a need for change. In turn, recognition of this is felt to be a necessary stage in the process of behavioural change (Day et al., 2007). Day et al. (2007) also noted that experiential interventions such as DbTs can help individuals with a criminal history to form an awareness of their problematic behaviours which will increase their readiness to change.

Furthermore, the theme of *hope* was derived from the synthesis of studies in this review. Participants reported that they were able to think positively about their future. Although belief in their ability to change was not directly measured, it could be inferred that an increase in hope demonstrated that participants had developed a belief in their own ability to change. As noted in readiness to change literature (e.g., Prochaska & DiClemente, 1983; Ward et al., 2004) this is a primary step in the process of behavioural change. This finding is also supported by Burrowes and Needs (2009) who state that "practitioners should seek to facilitate a sense of agency and help individuals recognize their personal responsibility for their future situation" (p.3). Also, in accordance with theoretical models around the process of change, it is noted that a lack of mental well-being can be a barrier to behavioural change occurring (Ward et al., 2004). As such, it is suggested that DbTs may be instrumental in the process of a positive behavioural change occurring.

DbTs in the current review which aimed to address feelings of anger and aggression were found to reduce such feelings: Studies 1 (Blacker et al., 2008), 4 (Harkins et al., 2011), 5 (Keulen-De Vos et al., 2017), 6 (Koiv & Kaudne, 2015), 8 (Reiss et al., 1998) and 10 (Stephenson & Watson, 2018). As mentioned, it is not possible to know whether these effects are long lasting or whether changes in scores on psychometric measures would translate to fewer acts of violence. However, the findings were promising. A further skill reported as having been acquired through DbTs is that of communicating with others. This finding was concurrent with the findings of the impact of other arts based therapies (Batcup, 2013;

Heskell, 1960; Maas, 1966; Philips, 1997; Seibel, 2008). Although none of the studies looked at recidivism as an outcome measure, it is evident that participants gained skills (e.g., pro-social communication) which could potentially help them lead pro-social lives. This claim is supported by Ward and Brown (2006) who state that "the best way to lower offending recidivism rates is to equip individuals with the tools to live more fulfilling lives rather than to simply develop increasingly sophisticated risk management measures and strategies" (p. 244).

Although not reported in the results section as it does not directly relate to the research question, it is of note that participants in the study by Stephenson and Watson (2018) made mention of their views of the Actor Practitioners (APs) who facilitated the DbT. In short, it was felt that the APs were trustworthy and respected them as individuals. This appeared to be a key factor in the success of the DbT. With reference to therapists' attitudes towards offenders, Ward and Brown (2006) highlight the importance of therapists being non-judgemental and acting with respect towards them, as "at the end of the day, most offenders have more in common with us than not, and like the rest of humanity have needs to be loved, valued, to function competently, and to be part of a community" (p. 244).

Lastly, whilst specific, in-depth details regarding the DbTs are not provided in the included papers, it can be said that DbT is an interactive process, typically using role-plays and allowing participants the opportunity to practice their skills in a safe place. Of particular note is the use of masks as a metaphor by the Geese Theatre Company. Geese "work with the notion that everyone wears 'masks', some as a habitual coping strategy, others more consciously…This technique enables audiences to explore the distinctions between external presenting behaviours and internal experience and can prove to be incredibly motivating for audiences and group members" (Geese Theatre Company: The use of masks, n.d.). Similarly, The DbT evaluated by Stahler (2007) used movement with masks to

encourage participants to physically express their inner world through a mirroring process. Masks were used to explore their addictive behaviours and were further used to act out participants' hopes and dreams for the future.

Limitations of the articles

Many of the studies reviewed in this review used one-sample experimental designs, which can compromise the internal validity of the study; it is not possible to determine cause and effect in such studies. In addition, as noted by Blacker et al. (2008), a variable which could have an impact on the reported effects are that participants may have been participating in other interventions at the same time as the DbT so it may not be possible to attribute changes in scores on psychometric measures directly to the DbT. For example, Colquhoun et al. (2018) delivered a three day DbT workshop to patient's concurrently engaging in a sex offender programme.

Furthermore, although pre and post intervention evaluations are commonly used as a method of measuring the impact of a given programme, it is noted that responses post programme may be impacted by the participant's memory of completing the measure prior to the DbT. However, researchers may justify using this approach for exploratory studies measuring shorter term impact of programmes. Harkins et al. (2011) however, report in the limitations section of their paper, that retesting participants using the same measures at pre, post and follow up timescales of an intervention may impact on post intervention outcomes One way to test the reliability of test retest psychometric scoring at different time points is a method called the Limits of Agreement (LOA; Bland & Altman, 1999). This process involves looking at the nature of the scores over the repeated measurements timescale. LOA plots the data to help interpret the agreement in scores at the different time points.

In addition, as noted by Blacker et al. (2008) and Reiss et al. (1998), positive changes found on self-reported measures do not necessarily mean that changes in violent behaviour will follow. Stephenson and Watson (2018) noted that a better measure of change in behaviour would have been to look at reported incidents of violence in participant's pre- and post-intervention.

The majority of the studies in this review used self-report measures (whether qualitative or quantitative), and as such findings must be considered in light of the potential for response bias (e.g., social desirability as highlighted by Blacker et al., 2008). Furthermore, none of the studies included a follow-up measure to measure whether the effects of the intervention had been sustained over a longer period of time.

Reiss et al. (1998) noted that drawing conclusions regarding the impact of DbT on a particular type of participant was challenging due to the nature of having different projects with different types of offenders. Participants were patients with varying offending histories, and either a mental health diagnosis or personality disorder. Similarly, the DbT reported in the study by Colquhoun et al. (2018) included offenders with a wide range of issues (i.e., mental disorders, developmental difficulties, learning difficulties, and personality disorders).

With reference to measuring anger, Reiss et al. (1998) explained that the participants that produced high anger scores may often experience heightened feelings of anger, however they may suppress these feelings as opposed to lashing out physically or verbally. As such, it is difficult to draw concrete conclusions from the findings, i.e., to ascertain what the participants who do not express their anger took from the intervention. In addition, in the Reiss et al.'s (1998) paper there was a lack of ethnic diversity within samples meaning that findings cannot be generalised beyond white Caucasians.

Sampling bias may have occurred given that participation in the programmes was not mandatory. As such, participants who engaged with the DbT may be considered to have an

existing positive view of drama and/or have an existing desire to change, making it more likely that they would engage with and, therefore, benefit from the DbT (Baim et al., 2002).

Of the quantitative studies, the largest sample size was 76 (Harkins et al., 2011) and the smallest was nine (Keulen-De Vos et al., 2017). Having a small sample will reduce the power of a study (Field, 2017; Hackshaw, 2008) thus increasing the risk of a type 1 or type 2 error[3], so findings must be viewed with caution. Keulen-De Vos et al. (2017) reported that the non-significant results found for the anger-mode experimental session could have been due to a low sample size. However, Hackshaw (2008) noted that "there is nothing wrong with conducting well-designed small studies; they just need to be interpreted carefully…data from such studies should be used to design larger confirmatory studies" (p. 1143).

Strengths and Limitations of this Review

Despite efforts to ensure that search terms and inclusion/exclusion criteria used in the review were such that all relevant papers were included, some studies may have inadvertently been excluded at the search or application of criteria stages. Furthermore, it was only possible to include studies published in the English language which may have introduced bias (Gregoire et al., 1995). Gregoire et al. (1995) highlight that having clear exclusion criteria with detailed linguistic criteria may result in applicable references being lost than by not using such a system. A strength of the review is that multiple relevant databases were searched as well as a manual search was carried out of reference lists of relevant articles.

The range of methodologies and measures used by studies included in this review made the synthesis of results challenging. However, the inclusion of both qualitative and quantitative designs was a strength in that the review included in-depth findings as well as the findings of studies with larger sample sizes that used validated psychometric measures.

[3] A type I error (false-positive) is when the researcher discards a null hypothesis when it is believed that it is true in the population; a type II error (false-negative) is when the researcher neglects to discard a null hypothesis that is considered to be false in the population.

Furthermore, the quality assessment procedure was considered robust as approximately 30% of the included papers were assessed by a second reviewer and all studies were found to be of acceptable methodological quality for inclusion in this review.

As noted above, DbTs evaluated by papers in this review differed in content and delivery (Table 3). It is therefore problematic to draw concrete conclusions that could be generalised across all DbTs, however, it is possible to consider elements which are common across DbTs included in this review.

A limitation of the review is the breadth of offending histories of participants included in the studies. Due to the range across and within studies, it is not possible to draw concrete conclusions about the impact of DbTs on particular offending groups. However, despite differences in DbTs and in participant groups, findings were consistent across studies regarding the positive impact of DbTs. This consistency may promote confidence in the findings of the studies and the synthesis included in this review.

Directions for Future Research

Due to the methodological disparity of studies included in this review, findings need to be interpreted with caution as they are not generalisable. In order to gain a fuller understanding of the impact that DbTs have on individuals it is suggested that further studies be conducted that include additional relevant measures of cognitive and behavioural change. For example, self-report questionnaires, self-monitoring, and behavioural observations may be useful (Hollin & Bloxsom, 2007). More specifically, in order to better evaluate the impact of DbTs on anger, measures could include, for example, number of anger outbursts, behavioural reports, and a wider range of standardised psychometric measures that aim to measure varying dimensions of the construct of anger. Furthermore, reconviction rates could be used to further assess the effectiveness of DbTs in relation to reducing re-offending behaviour. In addition, and as highlighted by the differences in outcomes noted in the studies

by Blacker et al. (2008), Watson and Beech (2008), and Reiss et al. (1998), it would be beneficial for future studies to investigate the impact of DbT on long-term anger versus impulsive anger. It is suggested that longitudinal studies be conducted that use valid and reliable measures of behavioural change.

Lastly, studies included in this review aimed to evaluate DbTs using data from participants. However, it is felt that the literature base would benefit from studies that explore the views of those who design and facilitate DbTs for forensic populations. In doing so, it may be possible to gain insight into particular elements of a DbT that *work* for participants, and also to gain insight into the challenges faced in delivering DbTs.

Overall Conclusions

Ten studies were identified for inclusion in the current review and findings were synthesised. DbTs differed in terms of content and delivery, however, common elements such as the use of role-play and reflection were noted across the included studies. A range of qualitative and quantitative measures were used in the included studies. Half of the studies included in this review investigated the impact of a DbT on levels of anger. Other issues explicitly addressed by DbTs were communication and relationship skills, problem solving, coping strategies, and confidence. DbT has a particular focus on tailoring the content and delivery to the needs of individuals; thus, showing adaptability and flexibility in its approach. Overall, outcomes from the studies reviewed indicate that DbTs can have a positive impact on participants in forensic settings in terms of reductions in anger levels and improvements in mental well-being, hope, problem solving, insight into behaviour, and communication skills.

DbTs have an advantage over manualised OBPs in that they do not require participants to have a certain level of literacy, they are less structured and are therefore able to be tailored to the needs of the individuals, and they use techniques which have been found to *work* in rehabilitation such as realistic role play (Andrews & Bonta, 2017) and reflection

(Mullen, 1999). The findings of studies in the current review are reflective of the findings of the positive impact of DbTs with non-forensic populations (Feniger-Schaal & Orkibi, 2020). Further robust research is crucial in order to advance knowledge and understanding of the use and effectiveness of DbTs within forensic populations. In turn, a robust evidence base for the use of DbTs should be established, which may lead to DbTs becoming more widely available to individuals in forensic settings.

CHAPTER 3

The Geese Theatre Company (GTC): An exploration into Actor Practitioners' (APs) views and experiences of the use of Applied Theatre with a forensic population

Abstract

Background

Findings have been encouraging regarding the effectiveness and positive impact of Drama based therapy (DbT) programmes within forensic populations in terms of reducing negative behaviours. However, less is known about the reasons underlying the positive impact that such programmes may have.

Objective

The aim of the current study was to explore the views and experiences of both former and current Geese Theatre Company (GTC) Actor Practitioners (APs) who deliver Drama based therapy (DbT) to individuals in forensic settings. APs' views were explored to gain insight into aspects of the programmes they felt were beneficial to clients, techniques they used to engage clients, the therapeutic relationship, training they had received, and challenges they had experienced in their role.

Method

Seven Geese Theatre Company (GTC) APs (four former and three current APs) took part in semi-structured interviews. All participants had completed the same six-month GTC mandatory training upon starting their employment. Interviews were conducted over the telephone. Interview transcripts were subjected to Thematic Analysis, i.e., codes, themes and sub-themes were generated from the data.

Results

Four key themes were identified: *The client journey*; *Actor Practitioner qualities and skills; Creativity vs. manualised interventions; Challenges and experiences.* These overarching themes were further broken down into sub-themes. The results are concurrent with previous findings in respect of the positive impact that GTC interventions can have. In

addition, the results add to the small body of literature around the techniques and approaches of DbT used within forensic populations.

Conclusions

The outcomes of the current study are discussed with reference to previous literature in the field of DbT and rehabilitative literature more broadly. In addition, implications for practice are provided with the hope that findings can inform other creative Arts Practitioners working with forensic populations and, more broadly, could add to the evidence base regarding *what works* in rehabilitative interventions.

Introduction

The Geese Theatre Company

Drama therapy and psychodrama techniques have been utilised in the treatment of adults, for example, those who have experienced trauma and/or have mental health needs (see Chapter 1 for further details). In addition, such techniques have been used with forensic populations by a range of organisations and individuals. One such organisation is the Geese Theatre Company (GTC), founded by Bergman (a drama therapist) in the USA in 1980 and then established in the UK in 1987. Initially Bergman's work was in the area of trauma. He proposed that drama therapy could be used for channelling anger and aggression. In reference to the origins of the GTC, Bergman (2017) stated, "We not only made original theatre pieces with the men and women inside these institutions, but we created new techniques to match the social and eventually psychological problems we encountered inside" (p. 324).

The GTC works primarily with prison populations, however, they also use their specialised skills to support people who have committed an offence and have mental health needs, necessitating detention under the Mental Health Act. The theatre based therapeutic model used by the GTC has also been used in community settings for people that have been involved with the criminal justice system, or who are at risk of becoming involved with the criminal justice system. Their interventions have also been tailored for use with individuals in the social welfare system, such as people in recovery from substance misuse (Watson, 2009).

Typically, GTC interventions aim to increase well-being, address mental health problems, and reduce anger/aggression. Key principles of GTC DbTs include using theatre to explore behavioural issues and using technical tools and methodology to instigate change. Essentially, at the heart of GTC DbTs are three key concepts: change, choice, and personal responsibility (Baim et al., 2002).

GTC projects/programmes are grounded in Cognitive Behavioural Theory, Role Theory, and Social Learning Theory, and are designed to help the participant explore and resolve both personal (inner) and social (outer) conflicts; thus, promoting positive change (Baim et al., 2002). Depending on the needs of the participants, tailored projects aim to address problem solving and communication skills; increase confidence and self-esteem; and assist in goal setting. The projects include, for example, observing performances by GTC Actor Practitioners (APs), storytelling, movement, and role-playing using masks (Baim et al., 2002; Harkins et al., 2011). In GTC DbT, masks are used as a prop to symbolise the social roles people play, and the underlying processes that underpin understanding of these social roles. There are a range of masks used in the programmes, each of which represents a persona, in other words, a role or attitude that the participant can relate to. The masks are assigned names: The Fist, Mr Cool, Good Guy, The Brick Wall, The Mouth, The Joker, Rescuer, and Poor Me. Participants who take part in an intervention are encouraged to consider which mask or masks they wear. They are then encouraged to (metaphorically speaking) take down their masks and to assist others in also removing theirs.

DbTs are used in criminal justice settings by the GTC in order to promote physical expression of emotions as opposed to trying to implement a strict dramatic protocol such as a scripted play (Smeijsters & Cleven, 2006). They use methods of theatre practice (e.g., role-play) to channel choice, responsibility and change within a physical space. The GTC employ APs who are trained and experienced in using evidence-based techniques with participants. They highlight that their APs are not psychodrama therapists and are not delivering psychodrama interventions (Baim et al., 2002). The GTC make a distinction between psychotherapy and the work they do. They highlight that the term Drama therapy should not be confused with Psychodrama nor Psychotherapy. They further stipulate that their approach is an intervention and not a treatment (Baim et al., 2002; Harkins et al., 2011).

Research Regarding the Impact of Geese Theatre Company Interventions

Multiple studies have been conducted in order to explore the impact of a range of GTC projects; the results of which are promising. For example, a UK service evaluation of a GTC programme (*Scratching the Surface*) for females in a prison setting was conducted (Stephenson & Watson, 2018). The evaluation investigated changes in levels of hope (using the Beck Hopelessness Scale; Beck et al.,1974) and mental well-being (using The Warwick-Edinburgh Mental Well-Being Scale; Tennant et al., 2006) prior to and following completion of the programme. The authors noted various limitations of their study (e.g., small sample size) but findings pointed towards the programme resulting in significantly increased feelings of hope as well as improvements in mental well-being in the women. However, the authors of the paper highlighted that the results should be treated with caution and that further studies are needed in order to draw robust conclusions (Stephenson & Watson, 2018). Their evaluation further included written feedback from participants regarding their views on the intervention. The responses revealed that they felt they could trust the APs and that APs genuinely cared about their needs. Further comments included that the mask exercises helped them to recognise the masks others wear, and that by understanding this they would be more equipped to manage social interactions. One participant in the study by Stephenson and Watson (2018) stated:

> It has made me see a better me and it's showed me new skills that I can use in everyday situations, and it will allow me to build up self-esteem, to have more confidence in myself and ability to carry things out that I find difficult (p. 20).

An earlier study evaluated the GTC's *Re-Connect programme* (a three-day intensive programme) with a sample of UK offenders, aimed at addressing factors associated with release and rehabilitation back into the community (Harkins et al., 2011). The programme

works on helping participants learn and develop tools to better manage social interactions and emotional arousal. It addresses issues of anger, aggression, and motivation to change, using a variety of methods such as theatre performance, role-play exercises, and masks. Levels of aggression, confidence, self-control and desire to change were assessed at the beginning and at the end of the intervention, using a psychometric measure of anger. Improved behaviour and engagement were observed over the three-day duration of the programme. Interviews conducted following the intervention highlighted that participants found the mask work useful and role-playing possible situations they might encounter following release into the community, such as attending a job interview, was reported to be helpful. Some participants also commented that they felt more confident following the programme and better prepared to deal with the future. Overall, these outcomes of the intervention provided some support for the short-term effectiveness of the programme.

Furthermore, Blacker et al. (2008) investigated the impact of a GTC extensive nine-day intervention with male offenders, which incorporated active exercises, role-play, and theatrical practices such as the mask. Exercises involved, for example, developing hypothetical characters in situations which participants could relate to, using self-reflection to understand anger habits and develop victim empathy. The programme incorporated the notion of choice, problem-solving, positive self-talk, and anger management techniques. The aim of the exercises was for participants to gain experience of using adaptive ways of dealing with high risk situations, such as the facilitators acting out an anger encounter. Pre- and post-intervention measures revealed a reduction in anger and aggressive feelings. Further findings revealed a learned understanding of how core beliefs impact on behaviour. Blacker et al. (2008) also noted that the programme is accessible for individuals who may have possible language and literacy deficits and who would, therefore, find more generic anger management programmes challenging.

The Role of an Actor Practitioner (AP) at the Geese Theatre Company

Initially in a group setting, the AP works to build an environment of trust, and to encourage an appreciation for everyone in the group. An AP works to promote the three key concepts previously mentioned: change, choice, and personal responsibility (Baim et al., 2002).

The *change* concept refers to the belief that everyone has the ability to change, that is, traits and behaviours are not static but dynamic (Baim et al., 2002). An underpinning thought is instilling in participants that they may not be able to change wider matters outside of their control, however, they can change things about themselves (Baim et al., 2002). This is explained through an example: if somebody committed a crime some years ago, they are not defined by their crime, rather they are considered to have the *power to change*.

The concept of *choice* refers to the notion that everyone has choices in life. It is recognised that people in secure settings are restricted in the choices they can make, and in addition, may have experienced an upbringing and home environment of similar nature. As such, the AP works to create an awareness of the choices that participants have at a micro level, such as choosing to be at a session on a given day, choosing to make eye contact with others in the group, and choosing how to respond to matters outside of their control. Furthermore, at a macro level (wider choices), APs aim to make participants aware of the opportunities they have to make choices concerning their future. In short, the AP uses their skills to increase a participant's understanding of the choices they have and the potential consequences of such choices.

The concept of *personal responsibility* advocates that if one fails to accept ownership of their actions then they are unlikely to change. This facet of the intervention focuses on encouraging people to take ownership of their behaviour. It is acknowledged that often criminal behaviours may have occurred in response to an environmental factor, such as an

unstable home. Participants are encouraged by the APs to move forward in life; they are encouraged to take responsibility for their actions as it is suggested that this will result in a sense of personal freedom. The GTC propose that for change to take place it has to come from the person; thus, they are the only person who can make the change. The AP encourages participants to own their identity; explaining to participants that identity plays a big part in theatre in terms of how people want to be viewed by others. The overall aim of the programme is positive change occurring in each individual (Baim et al., 2002).

Research on Practitioner Views and Experiences of the use of Drama Based Therapy (DbT) within Vulnerable Populations

In addition to soliciting feedback from participants who take part in drama/theatre based interventions, there have been a small number of studies which have sought to elicit feedback from practitioners who deliver the interventions to vulnerable populations. It is felt that by exploring the views and experiences of practitioners, those in the field can gain insight regarding the delivery and impact of DbTs.

Papagiannaki and Shinebourne (2016) explored the views and experiences of active creative art psychotherapists (art and drama therapists) who delivered a programme on the topic of self-stigmatisation to psychiatric patients in the UK. Six therapists were interviewed, and the data were analysed using Interpretative Phenomenological Analysis (IPA; Smith & Osborn, 2003). The study found that therapists believed that engagement with creative art therapy can break down issues related to self-stigmatisation. Self-stigmatisation refers to feeling inadequate due to being associated with undesirable social phenomena (Leff & Warner, 2006). Therapists reported that participants tended to have low levels of confidence in themselves, especially in the initial engagement stage. They further reported that participants tended to emphasise what they felt unable to do rather than what they could do.

Ultimately, their view was that the programme had a positive impact on those who engaged, and that art therapies can alleviate patients' feelings of self-stigmatisation.

Cassidy et al. (2017) explored key therapeutic components of change based on the experiences of both drama therapists (DTs) and drama therapy participants. Cassidy et al. (2017) highlighted the "importance for therapists to understand the processes experienced by the participant as a way of enhancing empathy, collaboration and attunement" (p. 2). The researchers applied a Grounded Theory approach (Glaser & Strauss, 1967) to their data to collate the following key themes: the physicality of experimental self-play; the space to be able to explore; and experimentation for potential change. A common theme across feedback from DTs was the importance of taking time to build a therapeutic alliance with the participant; only then can a participant feel safe, which, in turn, leads to participants feeling liberated and able to properly explore avenues for change. These findings led the researchers to conclude that this therapeutic approach facilitates an exploration of a common language, specifically, to communicate through physical expression.

Furthermore, Keisari et al. (2018) investigated Drama Therapists' (DTs) perspectives on delivering a Playback Theatre (PT) therapy for older adults. The authors described PT as a creative process by putting one's *life-story* at the heart of the improvisational theatre intervention. A total of eight therapists were interviewed regarding, for example, their thoughts on being a conductor of groups (e.g., training and which groups they work with); the therapeutic process (e.g., degree of cooperation, challenges, guiding principles); life stories of older adults that had emerged; and participant attitudes towards acting and improvisation. The collected information was subjected to a Grounded Theory approach to draw out themes. The data showed DTs reported the following: PT encourages people to feel free to be playful and from this they are able to be more self-reflective; being open through play encourages people to be comfortable within a group, and encourages people to dip in and out of others'

stories; PT allows people to reminisce of earlier times in their lives, which they felt was therapeutic for participants; and PT allowed participants to reflect on their lives which they described as an awakening of the self. In addition, participants gave a broad overview of a variety of methods they used to ignite participants' memories to tell their life-stories. They reported on, for example, how life stories could be reconstructed and brought to life, and how the theatre was viewed as "a healing space" (p. 75).

Rationale and aims of the current study

As outlined above, studies have looked at the impact of GTC interventions with offender populations in terms of reducing negative behaviours (e.g., aggression) and issues relating to mood (e.g., depression) of which, the results are encouraging (Blacker et al., 2008; Keisari et al., 2018; Stephenson & Watson, 2018; Papagiannaki & Shinebourne, 2016). In addition, as outlined above, research has been conducted which has explored the views and experiences of practitioners who use drama/theatre as a form of therapy. Insights provided by such practitioners can be of use in developing an understanding as to how drama/theatre interventions can be applied to vulnerable populations. However, to date, no study has explored the views and experiences of those who deliver GTC DbTs to UK forensic populations.

As mentioned above, the GTC are unique in the ways in which they utilise theatre/drama in order to help vulnerable adults (e.g., offender populations). As such, it is suggested that it would be beneficial to explore the views and experiences of APs in order to gain insight into, for example, the impact they feel their inventions have on participants, and what it is about the interventions that they perceive to be effective. It is felt that such information could inform other creative arts practitioners working with forensic populations and, more broadly, could add to the evidence base regarding *what works* in rehabilitation with reference to non-accredited interventions. In short, the current study aimed to explore the

views and experiences of both former and current GTC APs who have delivered DbT in the criminal justice system. The views and experiences include, but are not limited to, APs' thoughts on the process of change within a participant during the intervention and the therapeutic relationship.

Method

Research Design

A qualitative approach was deemed to be the most appropriate method in addressing the research aims, for example, to explore participants' views and experiences of their current or former work as an AP. The current study deemed semi-structured interviews to be the most appropriate, as they are viewed as a flexible means by which to gather data and prompts can be added to encourage participants to expand on answers given (Robson, 2002). It was further deemed appropriate to record interviews on a Dictaphone in order to provide a full and reliable record of participant responses. In addition, research has highlighted that by the researcher recording interviews, they are able to be more interactive with the participant than if it were necessary to take notes, allowing the researcher to be more responsive during the interview (Loftland et al., 2006).

Semi-Structured Interview

All questions were open-ended, the aim being for participants to use free recall when expressing views and experiences. Prompts were in place to obtain more information, if needed, to add to the initial responses provided. Braun and Clark (2013) champion this approach; they explain that the interview flow is guided by participant responses in order to minimise the potential for researcher bias. Questions were formulated from the research aims to explore: APs views and experiences of their role; reflections on the GTC training received; aspects of the programmes that are/were the most memorable for them; aspects they felt had

the greatest impact on clients; interpersonal skills they believed are/were key to client engagement with the intervention. Questions included, for example: *"*Can you tell me your thoughts on elements of the intervention (if any) you believe to be more successful to the participants?"; "Can you tell me about the types of interpersonal skills required from participants, in your opinion, for successful engagement with the intervention?" (See Appendix I, for full interview schedule).

Sample

The data collected for this study were from former (n = 4) and current (n = 3) GTC APs, located in the UK (n = 5) and Australia (n = 2). The total number of participants included in the study was seven. All participants had completed the same six-month GTC mandatory training upon starting their employment. All participants had worked as an AP for the GTC for at least two years, with the longest serving former employee having worked there for over 10 years. The two current employees had worked there for at least 10 years on a flexible basis, taking on work as and when needed. The sample comprised of four males and three females. It is of note that the original research proposal was to interview at least six current GTC APs to provide approximately equal numbers of former and current employees. However, due to Covid-19, the majority of current staff were furloughed, and it was not possible for them to participate in the project resulting in a smaller than desired sample size. The amendment was approved at School and College level within the University of Birmingham. It is recognised that this was a limitation of the present study (this is discussed in the limitations section at the end of the chapter).

Procedure

Initially, the director of the GTC was contacted to discuss whether he felt the project would be of value to the field and whether he felt that both current and former APs would be

interested in taking part in the study. Following his positive response, ethical approval was sought and gained from the University of Birmingham Ethics Committee (ERN_19-1965). Potential participants were sent an initial engagement letter via the Director of the GTC (see Appendix F), which explained the aims of the study. Once they had responded to this letter, they were sent a comprehensive participant information sheet providing more information regarding the stages of the study (Appendix J). Following expression of interest to participate in the proposed study, the Director of the GTC provided the researcher with a list of participants along with the contact details of those that agreed to be contacted directly by the researcher. An introductory email was then sent to all participants (Appendix F) and arrangements were made for the interviews to take place.

Interviews were conducted over the telephone for convenience and due to Covid-19 social distancing regulations. Conducting interviews over the telephone has historically been considered to be inferior to face-to-face interviews in that it is not possible to use non-verbal communication and it is considered to be more difficult to build a rapport (Lavrakas, 2008). However, there is no evidence to support this assertion and it has been suggested that conducting an interview over the telephone may result in the participant feeling more relaxed and better able to disclose sensitive information (Novick, 2008).

Participants were asked to ensure they were in a quiet and private space for the duration of the interview. The researcher read aloud all items on the consent form (Appendix E) and asked the participants to state whether they verbally agreed to each one. At this stage they were reminded that the interview was being audio recorded. The researcher ensured participants understood the aims of the study and what they were consenting to. The structure and expectations were further iterated to participants in terms of how long the interview may take. At this stage, the researcher gave the participants the opportunity to ask any questions before starting the interview and made them aware that they could withdraw from the study at

any point during the interview if they so wished. Interviews lasted between 45 and 65 minutes.

Following the interview, participants were asked if there was anything further they wanted to add to their responses that they had not already discussed and were also asked if they had any questions for the researcher. The participants were further reminded that they would need to contact the researcher within two weeks following their interview should they wish to withdraw from the study. Following this timescale, the data would have already been included in the analysis and therefore it would not be possible to retract individual information. Participants were sent a debrief sheet via e-mail which provided contact details should they want to receive a copy of the final report (Appendix H). Participants were thanked for their time and received remuneration of £50 for their time, which was funded by the Birmingham University.

Ethical Considerations

Ethical approval for the study was sought and granted from the University of Birmingham Ethical Review Committee in March 2020 (ERN_19-1965). The Committee queried whether there was any possibility that participants may provide identifying or particularly sensitive data about individual prisoners they had worked with when answering the interview questions. The information sheet was therefore reviewed to include a request that participants refrain from disclosing any information by which a client they worked with could be identified. However, in the event of disclosure of sensitive information, the researcher agreed to ensure that these details would not be included in the transcript (i.e., the information would be omitted), and the audio recording would be digitally erased following transcription.

Consent

An information sheet along with the consent form was emailed to all participants

(Appendices E and J). Due to the nature of how the interviews were conducted (i.e., on the telephone), all items on the consent were also read aloud where participants were able to verbally consent to each point prior to the interview commencing. Contact details of the researcher along with those of the researcher's supervisor were provided should any of the participants have questions or want to request to withdraw from the study. It was explained to the participants that they had up to two weeks from the date of the interview to withdraw from the study. None of the participants chose to withdraw from the study.

Anonymity and Confidentiality

The data was collected using a Dictaphone and temporarily stored on an encrypted USB device before it was uploaded to Research Data Store. This is a data storage service held by the University of Birmingham. It is a secure password protected system for researchers to store and access their data. Once consent for participants in the study was obtained, participants were also informed that verbatim quotes would be anonymised to be published in the final report. Each participant was given an ID code, which was non-identifiable. This was done to be able to identify individual audio recordings of the interview transcripts should any of the participants request to withdraw from the study. No personally identifiable information was present once data had been anonymised and data analysis commenced. Any information provided during the interview by which participants, or their clients could be identified was omitted during transcription. All participants were advised that their audio data would be deleted from the USB and the research data store (including digital copies of consent forms) following transcription, held for up to two months for the purpose of analysis and write up.

Risks

Participants were asked to refrain from disclosing any personal information regarding clients they had worked with. If this had occurred, the protocol would have been to omit this information from the transcript prior to the audio recording being deleted following

transcription. Furthermore, a debrief form (See Appendix H) was emailed to all participants following their interviews. This document further explained the aims of the project and provided details of the researcher and the research supervisor in the event anyone wanted to withdraw their data from the research. In addition, it provided a further opportunity for any issues or queries to be raised.

Analysis of Data

The data was analysed using Thematic Analysis (TA), a qualitative data analysis approach used to explore patterns of data (Braun & Clarke, 2006). The methodology involves coding data to develop themes. Thematic Analysis is considered a comprehensive and adaptive approach in organising and making sense of information. Conducting TA will vary from one researcher to the next depending on the particular research aim and topic (Braun & Clarke, 2006).

Following discussion with the academic supervisor, the researcher deemed TA to be the most appropriate method to use with the current data set due to its deductive approach, which allows for some pre-existing codes to be applied to the transcript (a-priori themes/sub-themes). For example, Braun and Clarke (2013) describe three sub-categories of Thematic Analysis, which informed data analysis in the present study. Firstly, a *realist* system which looks at experiences, meanings and reality. Secondly, a *constructionist* system, which looks at the ways events and meanings all applied within a society. Thirdly, a *contextualist* system which looks at elements that involve both reality and societal notions (Braun & Clarke, 2006). Due to the nature of the current study (i.e., views and experiences of APs delivering creative programmes within criminal justice settings), the contextualist system of TA was deemed appropriate (i.e., looking at meaning and experiences of the APs whilst acknowledging societal influences). For a theme to be established, a congruent pattern of responses within the data set was clustered (Braun & Clarke, 2006).

In line with Braun and Clarke's (2013) guidelines, the researcher applied two methods of theme development: an inductive (*bottom-up)* technique and deductive (*top-down)* technique (Braun & Clarke, 2006). The inductive approach focuses on a data-driven understanding whereby the researcher refrains from trying to fit the data into a pre-developed coding framework (Braun & Clarke, 2006). A deductive technique is the opposite as the researcher's interest in the topic area drives the analysis. The latter technique, however, can result in a lack of rich description of the data overall. Once considered, the researcher and the academic supervisor agreed on using both inductive and deductive techniques when analysing the interview transcripts.

In terms of types of themes, this is defined in two spheres: *semantic* and *latent* (Braun & Clarke, 2006). Semantic relates to the specific meaning of the data, as opposed to latent, which relates to going deeper, past the overt meaning of the data (i.e., underlying concepts) (Braun & Clarke, 2006). For the purpose of the current study, a focus was placed on semantic meaning rather than a more latent, interpretative approach. Data analysis is considered fluid as opposed to a linear process, and as such the researcher can revisit data analysis at any stage. Before theme development, the researcher familiarises themselves with the data by re-reading transcripts and developing codes. All data was considered line by line, and codes were assigned which represented features of the data. Assigning codes involved "a continual bending back on oneself – questioning and querying the assumptions we are making in interpreting and coding the data" (Braun & Clarke, 2019, p. 594). Codes were then grouped into themes and sub-themes.

Interviews were transcribed verbatim to ensure all spoken words were considered. To eliminate researcher bias, a fellow doctorate student from the University of Birmingham reviewed a small random selection of excerpts from three interview transcripts to establish some inter-rater reliability. Potential coding labels for themes and sub-themes were fed back

to the researcher and duplication in coding descriptions were unified. In summary, positive agreement was achieved for codes within the data. This process helped the researcher ensure themes and sub-themes were data-driven, thus minimising researcher bias. In addition, themes and sub-themes were discussed with the researcher's academic supervisor, and some subsequent changes were agreed on. In accordance with Braun and Clarke's (2019) paper, reflexive practice is considered to be an essential component of the TA process.

Results

This section presents themes identified from the data. Quotes from the transcripts are presented with participant identifiers P1 through to P7.

Table 4
Themes and sub-themes derived from Thematic Analysis of interview transcripts

Theme	1. The client journey	2. Actor Practitioner qualities and skills	3. Creativity vs manualised interventions	4. Challenges and support
Sub-themes	1.1 Pre-intervention attitudes 1.2 A change for the better 1.3 Group dynamics	2.1 Why GTC? 2.1.1 Training and experience 2.1.2 Helping others 2.1.3 Understanding and empathy 2.1.4 Curiosity for criminality 2.2 Communication is key 2.2.1 Techniques 2.2.2 Building a rapport	3.1 Flexibility 3.2 Something different 3.3 Masks	4.1 Challenges 4.2 The benefits of challenging yourself 4.3 Training 4.4 Learning 'on the job'

Theme 1: The Client Journey

APs were asked questions relating to the types of attitudes they felt clients held when they started the programme. They went on to discuss the ways in which they felt clients'

attitudes changed over the course of the GTC intervention. They also spoke of how the interventions helped clients reflect on their lives and behaviours and felt that these reflections and insights resulted in positive outcomes. In addition, participants discussed the impact of group dynamics on the process of change.

1.1 Pre-intervention Attitudes

In addition to attitudes and behaviours that clients may have displayed prior to their time in prison, APs believed that being in restrictive environments such as prisons or secure hospitals, could lead to clients having additional complex issues that need to be addressed prior to being able and/or willing to engage with the intervention. They reported that breaking *the norm* of behaviour in prison can be very difficult for clients. APs explained that often clients who attended had the attitude of, "oh ok, I'll watch what they do, I can kill two hours" (P2).

Participants reported that, in some cases, clients did not appear to want to be in the intervention, and that they were perhaps attending as they felt it was preferable to the alternative of daily prison life. For example, one participant commented, "might as well, got nothing better to do…." (P4). It was noted that sometimes clients arrived in the room with their guard up. It was suggested that this could be due to a lack of trust in the APs and sometimes not knowing the others in the group. Furthermore, it was noted that clients may come to the first session displaying behaviours, which may serve as barriers to engagement. For example, one AP discussed working with young males who were at the beginning of their life sentences, describing them as "a group of tough guys, cocky and acting very jack the lad like" (P7).

APs noted that, in their experience, the majority of clients who attended were pessimistic regarding the potential benefits of the interventions. However, APs also described

how some clients chose to attend the intervention as they were curious as to what it would involve and wanted to try something new. One AP was of the belief that "there needs to be an element of bravery [of the client]" (P5), and another reported that it is beneficial if a client possesses "…the ability to try new stuff, to take part, a willingness to give it a go" (P1). Similar suggestions by APs included, for example, that "…there needs to be an initial baseline – a desire for one to shift things, a willingness to change…" (P2) and that "confidence helps with engagement" (P2).

Conversely, one AP noted that "the ones that don't initially engage, get the most out of it" (P6). Others also reported instances where clients did not want to engage at first but who had breakthroughs, indicating that, overall, APs were of the opinion that GTC interventions could be beneficial even when clients did not express an interest at first.

Furthermore, another AP reported feeling that, despite barriers that an individual might display, the fact they sign up for the intervention shows that they have some hope that it may benefit them.

1.2 A Change for the Better

After discussing the attitudes held by clients on arrival to the intervention, which included reports that some clients were guarded and put up barriers, APs described their views on subsequent changes, if any, in the clients' attitudes and/or behaviours. APs spoke about such changes and provided examples of particularly memorable client cases.

APs described how, through the use of certain techniques/traits (see section 3.3), clients' guards and barriers would come down. One AP described this as, "…little by little, walls come down…" (P7). It was reported that such guarding is reduced after clients get to know others in the group and begin to trust APs. A focus is placed on this at the beginning of

each intervention as “ice breaker exercises help to get to know people and bring their guard down” (P7).

Some examples were provided of instances where APs recognised that a client had lowered their guard: “One of the guys cried in the group in front of the others…it was an opportunity for letting their guard down when they usually have so much front” (P2), and another AP described how a client who to begin with had been very closed off, had shared with the group that he had been abused as a child: “…by the end of the session he shared his story freely, and his jacket came off. It felt like a weight had been lifted, talking about this in the right way with the right kind of people” (P2).

APs also noted other changes in clients. The use of techniques such as masks (see section 3.3), were found to help clients reflect on their lives and behaviours, leading to observable changes in them. An AP gave an example of one particular case:

> We slowly challenged his defence mechanisms and barriers. We used fragment masks to do this. When using the ‘fist’ fragment mask it was like the penny had dropped. He kept saying how stunned he was and that it really made him think. He took on the mime scene within the programme and was great at it. He said that he hadn’t spoken to his partner for years but came to the second session telling us that he had contacted her to say sorry and told her he wanted to be in touch more. He said – ‘I wanted to lift my mask and tell her what’s going on’. At the end of the programme, the guy took one of the starring acting roles in the end of programme performance, and he said after it that it was like doing drugs. We were later told that this guy’s medication had been reduced and attributed it to the Geese work he had done. (P2)

Other APs also reported witnessing increased self-esteem and confidence in participants following their engagement with an intervention. They felt that this confidence led to a willingness of clients to try new things, including revisiting hobbies and

relationships. APs reported that clients learnt to be open and honest and to self-reflect. Furthermore, one AP reported that completing a GTC intervention can help clients to take some control and say *no* when necessary.

It was noted that clients became able to step outside of their comfort zone. Furthermore, one AP commented that GTC interventions help clients "to think outside of the box, to think outside one's own bubble" (P6). APs further reported that clients became more tolerant of others by identifying, understanding and empathising with what others had been through, leading to increased respect for others. It was reported that, in part, this was achieved by clients being encouraged to interact more and in different ways with others by talking, listening, and breaking down social norms. It was reported that, as a result, "…they smile more and become a lot lighter" (P2) (see section 1.3 for further detail on the experiences of group dynamics).

One AP commented that GTC interventions "create opportunities for pressing the reset button" (P2). With particular reference to the game element of the intervention, another AP noted that it "…challenges things – it's that that makes people question things and make a different choice, allowing them to see other opinions, thinking of themselves outside of prisoner" (P6). APs reported that clients may engage with the prison regime more openly during and following a GTC intervention and that they may be better able to form friendships with fellow prisoners. It was also suggested that being exposed to the arts may encourage clients to do more of it when released from prison, if the opportunity was provided.

APs went on to report that taking part in a GTC intervention can provide a more accurate understanding of how they are perceived by others and can provide them with the opportunity to reflect on their own lives; thus, providing insight regarding their own behaviours and attitudes: "I think holistically – they are able to think about their own

thoughts, feelings, behaviours etc…, to start thinking about the reality of YOU can be hard but helpful" (P6).

Finally, it was noted that these reflections and insights can ultimately lead to positive changes. For example, one participant stated: "It can help to gain focus and perspective - where they want to be and what they want for their future, optimism and hope, moving towards a goal, and having the confidence that they can do it" (P1). Another AP suggested participants of the interventions "… can gain insight into themselves, they can look at things more objectively, symbolically and gain control over it" (P7).

1.3 Group Dynamics

Group dynamics (i.e., friendships, hierarchies) were reported as having an impact on individuals within the group. In sub-theme 1.2, the impact of communication within the group was touched upon in reference to change, however, APs made further comments specifically regarding communication within groups and, as such, this is reported as a separate sub-theme. The majority of comments regarding group dynamics were of a positive nature.

APs provided examples of the types of group dynamics they have witnessed in their role. It was reported that clients needed to be tolerant of others. For example, they need to allow others to talk, and they need to be willing to listen to what others have to say. They believed that clients need to feel safe enough to talk about experiences with the group and be respectful of others. They need to have the capacity to work collaboratively with others rather than dominate others:

> Often more dominant or powerful people in the group lead or initially dictate how others engage/respond, this can be both negative and positive. These people can be the key that unlocks the rest of the group, it's a status/social thing. The negative

points are that they fear they may be laughed at from other group members, especially in prison settings. Dominant group members can impact on others. (P7)

It was further noted that the interventions encourage collaboration between clients as they recognise the benefits of such collaboration. In this respect, P5 commented "…as the programmes are very pro collaborative stuff – working with others, sharing with others, make meaning together, they come to an understanding together". Furthermore, APs reported that being in a group can help clients see things from different perspectives, which can be beneficial for them.

One AP reported that taking part in group work "frees up communication" (P5) and believed that the experience of communicating with members of the group gives clients the skills and experience to communicate more effectively with people outside of the group; for example, they are able to "…nurture pro-social friendships with fellow inmates" (P1).

APs spoke of additional benefits for clients from interacting in a group setting: "Being in a group can create great bonds and connection with the people in the group, which can make people feel safer" (P4). It was felt that these bonds and connections can result, in part, from group members recognising similarities between themselves and others in the group. Groups were described as a safe place to share with others.

It was reported that the group provides "…a supportive environment and can break down social norms, gives them a sense of identity and community" (P5); and "They get to hear different perspectives and listen to others' opinions" (P6). One AP went on to explain that communication within the group can give clients insight regarding views held by their peers: "…some people may assume that others think the same as them; people realise there are other ways of thinking about stuff, and that's ok" (P6). It was felt that a further advantage of being in a group setting is that it can help people gain an understanding of how they are viewed by others, offering a "public perception of themselves" (P1).

Another reported benefit of the group context was that, through watching others and being encouraged to take part, people with less confidence will end up taking part in exercises. This, in turn, can help boost self-esteem and confidence. APs reported that watching others express themselves can help quieter group members speak out; helping them "to come out of their box" (P5).

More generally, APs discussed that being in group settings encourages people to talk, laugh and have fun with each other. An AP recalled that one group member made a comment about another group member, saying "I've never seen that guy smile before" (P7). In short, APs reported that a safe, supportive group environment benefitted clients in a range of ways as outlined above.

Theme 2: Actor Practitioner (AP) Qualities

APs were asked about the circumstances and/or background that led them to apply for employment with the GTC. Participants spoke of the appeal of working with people in prison/with criminal justice histories, and the qualities (professional and personal) they felt necessary to possess in order to be effective in their role.

2.1 Why GTC?

APs spoke about what led them to the role as well as what kept them in the role. Participants spoke of their past training and also of their desire to help others, their empathy for others, and their curiosity about individuals with a history of criminal offences. In the main, APs felt that it was a combination of these factors that made the role appealing to them.

2.1.1 Training and Experience. Many APs reported having a background in drama and having a degree level qualification in the performing arts. Some individuals noted that they were not aware of the existence of this type of work but were keen to know more when they heard of the work that GTC did: "I'd never heard of prison theatre, it was new to me,

looked fascinating" (P5). Some participants reported that they had heard of the GTC at drama school, with one stating that, "it made me aware of a company doing work I would love to do" (P6). Some APs commented that they were keen to gain employment in a role that required their skills and training as an actor: "I initially wanted to use it as a stepping-stone, I never thought I'd get a job actually using my degree"; "I wanted to use my theatre qualifications" (P1). A minority of APs felt that acting/theatre was their main skill, for example, "…the only thing I was good at was acting at school" (P2).

Interestingly, many reported that they had already worked with vulnerable groups in the context of support or care work such as young people, ex-offenders, those with mental health issues, homeless individuals, and people with substance misuse issues. Different experiences were reported with some saying they enjoyed support/care work, whereas others did not. One participant had prior experience of using theatre in the context of support work, saying: "I was working with kids that had been kicked out of school, they had behaviour problems. We worked with them using theatre, using theatre to look at addiction/their stories, and I loved it" (P4).

2.1.2 Helping Others. Some APs' comments centred on their desire to help others being a key factor in what had drawn them to the role: "I met a lot of arrogant actors, and I felt there was more that I could be doing – social change. I went on to study community issues, I wanted to use acting to explore these issues" (P5). One AP reported that when they were doing their drama degree, they went to work abroad with underprivileged children, which fuelled their desire to use their skills to help others:

…it changed what I wanted to do work wise, I used some drama with the kids there. I was excited by therapy theatre, I had done some acting after university, I wanted to do something more therapeutic…I wanted to be part of a process of change, being someone that people felt they could talk to /work with, to make a difference. (P1)

Further responses highlighted that APs wanted to help people by working with them to change their behaviour and enable clients to reach their potential. A further AP discussed how they worked on projects with people residing in hostels (i.e., ex-offenders with substance misuse issues). They used theatre to look at clients' stories regarding their addiction. This AP went on to state, "…that's how I wanted to use my theatre stuff" (P7). APs believed it to be a perfect combination of the two, in other words, the work the GTC does, brings "…theatre and rehab together" (P2).

Another AP went on to note that drama, in particular, was a powerful tool in encouraging positive change: "...being able to bring something new to people, making a difference, that's the power of drama has...I wanted to make human connections and do something meaningful" (P2). In short, APs noted that they wanted to be someone who used their skills to make a positive difference to people's lives.

2.1.3 Understanding and Empathy. Some APs made reference to their awareness of the circumstances clients had experienced, which resulted in them being imprisoned. In so doing, they displayed empathy for clients, putting themselves in their shoes. One AP made explicit reference to this, stating "…the strength of empathy – we have openness to connect and feeling others' pain" (P6).

Comments were also made which denoted a belief that, to some extent, the criminal actions of individual participants could have been as a result of external factors, thus displaying understanding and empathy for clients. For example, a few APs reported a belief that flaws in the criminal justice system and wider government can put some individuals at a disadvantage in life, impacting how people behave. Structural issues in society, which lead to poverty, were cited as being linked to criminality.

In addition, one AP (P3) reported that life can be complicated and suggested that had these individuals experienced less challenging life circumstances, they may have made

different, pro-social choices in life. It was also noted that people are complex and will behave differently when faced with conflict. There was a general recognition that many clients had lacked the positive opportunities afforded to others in society, suggesting that "…[they are] the underdogs of society, haven't had their voices heard, haven't had opportunities that others have had" (P7).

Another AP provided a personal response whereby they disclosed that they felt they may have ended up in prison had it not been for their involvement in drama: "I could have been in the same position if it wasn't for drama, criminal life…I can relate, life happens. I wanted to be part of helping turn their life around, make the right choices, give them hope" (P4).

2.1.4 Curiosity. A minority of APs mentioned that they had a general curiosity about crime and criminal behaviour: "I love acting, and I am also fascinated with real life/high profile crime"; "I have a macabre interest in true crime" (P4). One AP reported that, through their role with the GTC, you get to see into a world that you only read about in the media. For example, getting to meet high profile, notorious criminals in high secure prisons, stating that "it was an interesting way for me to make money" (P3).

2.2 Communication is Key

APs highlighted the need for good communication skills in their role. They discussed techniques that they used to connect with clients and, in addition, gave examples of ways in which they built a rapport with clients.

2.2.1 Techniques. Participants discussed their views on communication techniques they used with clients, which helped clients to engage with the interventions. For example, the importance of listening to clients was noted. Techniques such as active listening were reported to have been used, whereby, they focus on what a client is saying. In addition, it was

noted that emotional intelligence should be used when trying to determine when to step back and when to push forward in interactions with clients.

The importance of open body language was mentioned by a participant as it was felt that this can put clients at ease. There was a recognition that all clients are different as "some clients can be intimidating, others may be more guarded" (P4), which meant there is a need to get to know individual clients to break down barriers. It was felt that by asking and answering questions, it would be possible to find common ground in order to help break down barriers. One way to do this was during ice breaker exercises, which were noted to "…help also to get to know people and bring their guard down" (P2).

In chatting with clients during exercises in a session (and making the effort during break times) it was felt that it was possible to show clients that APs had a genuine interest in them: "Exercises help also to get to know people, bringing their guard down, showing them that you're normal…" (P1).

Having knowledge of a client was also noted as being helpful in creating performances that are more authentic (i.e., that reflect the client's experiences). One AP recalled that interventions often involved getting information from a client about their life and then using this in a performance: "Where you reflect the client's world by obtaining information about the area they are from/something they're interested in. I then put a bit of that into the performance which reflects their reality" (P6).

2.2.2 Building a Rapport. APs discussed ways in which they engaged with clients; providing specific examples of approaches as to how best to interact with clients on a personal level. APs reported that it was necessary to be approachable and friendly and to not try to put on an act. In the same vein, it was suggested that APs needed to be authentic, down to earth, and grounded in themselves as in so doing it would be easier to build a rapport with clients. It was also suggested that there is a need for it to be clear to clients that APs believed

in their own work, showing confidence and that they cared about their work. However, it was further noted that APs needed to not be afraid to get it wrong at times.

A common suggestion was that it is necessary to show respect for clients, that is, they should be treated as equals and as human beings and the stigma of their criminal behaviour should be dismissed. With equality in mind, it was also suggested that APs should not ask anything of a client that they would not be willing to do themselves. It should be recognised that clients are experts on themselves and their lives. It should also be appreciated that clients may well have experienced challenges in life and have a lot of life experience and this knowledge and experience should be respected. With this in mind, one AP highlighted the importance of not acting or speaking in a patronising manner with clients.

Finally, some APs reported that using humour was beneficial to rapport building and the effectiveness of the intervention. With one AP going so far as to say, "…when humour is at the core, you see real humility" (P3).

Theme 3: Creativity vs Manualised Interventions

APs were asked about what the GTC (and the DTP approach in general) could provide to the client group that other interventions (e.g., manualised interventions delivered in prisons and/or community settings) might not. They discussed their thoughts on the flexibility and freedom that they as practitioners (i.e., to tailor the delivery and content to the particular group or individual) and the clients (i.e., attendance is not mandatory) are afforded within the interventions. They also commented on the more general use of applied theatre techniques and how these differed from manualised interventions. Although there are multiple techniques used within GTC interventions, the key technique that was discussed with reference to its effectiveness was the use of masks, and as such, this is presented in a separate sub-theme.

3.1 Flexibility and Freedom

APs commented on the flexibility and freedom of DbTs as compared to other more uniform interventions that are often mandatory to attend for offenders. They highlighted that their interventions (and others like it) are attended on a voluntary basis so, from the outset, clients do not feel forced to attend and to engage with the APs; clients are made aware that they can leave at any time, which may lead to them being more likely to engage. However, as noted in theme 1, there can still be resistance to engage from some clients.

APs highlighted the flexible nature of DbTs in terms of being able to tailor their method, approach, and activity to the specific needs of the group; they are not constrained by a set manual or format as "DbTs allows you to use your own methods to engage" (P7). Due to this ability to be flexible in the approach they take, APs discussed how "it provides the flexibility to adapt to specific groups of clients" (P1) in terms of the content and pace; suggesting "the approach is people driven rather than standard stuff re: psychology, it's bespoke" (P6). It was noted that, unlike some other interventions, DbTs key learning outcomes are flexible from the outset and can change during the course of the intervention according to the needs of the group and individual participants.

APs spoke of being reactive to what clients say, for example, "We can ask them what this situation looks like, [and then] act it out…" (P2). APs highlighted that this approach means that clients can open up and APs can challenge things they say. Participants reported that clients can come to a realisation or make an independent decision, which is specific to their own pathway, rather than being told what to do, and therefore making things feel more *real* for an individual. Many APs commented on how arts based programmes are less conventional than manualised interventions in that they are not school-like because there is no classroom and no traditional teaching element.

In addition, it was noted that different people have different learning styles. Manualised interventions may rely on specific learning styles, whereas "creative arts programmes are also good for different learning styles. Creative arts are not prejudice; the arts cater for all" (P2). Other APs went further to state that, "Arts programmes are good for people who find it hard to engage in heavy programmes" (P5) and "…arts programmes are good for visual and active learners, and for people who have low attention spans" (P3).

3.2 "There's a Magic About Theatre"

APs were asked what they feel is different about the arts (in particular, the GTC) compared to other interventions clients may be offered. APs highlighted that their methods are designed to "…allow people to find their own way…" (P6), rather than to teach them per se. One AP simply stated that, "there's a magic about theatre" (P7) perhaps implying that theatre is a unique, special and powerful tool.

APs gave details regarding how unique and distinctive the interventions are and the impact that this can have on clients. For example, "watching the performances - people feel entertained/inspired" (P7), and "…they give opportunities to make personal connections which is something not necessarily found in other interventions" (P1). APs went on to comment that the creative approach is beneficial in a number of ways. For example:

> You can find a different identity if you associate yourself with an art form, look at yourself in a different way, which is important for reoffending – instead of thinking as themselves as a prisoner, they can also say I'm an artist or a runner for example, something else other than negative stuff, which is important for identity, to feel pride. (P2)

APs commented that DbTs offer a unique platform for clients to explore thoughts, feelings and behaviours in a creative way. The intervention allows APs and clients to create

characters and scenarios that clients can connect with. It may be, for example, that the characters they create are based on their own life experiences and, as such, are specific to them. For example, one AP commented, "they can go into their past/their life to create fictional characteristics and scenarios, it's a safe space" (P5). Another AP commented on the impact of the work, stating "the arts can reach a part of someone that no other invention can, it's about the soul, reaching and connecting with a deeper part of yourself, you can be someone else - lose yourself" (P3). Lastly, another AP stated, "…you can create something beautiful or ugly, making meaning for yourself and the people around you" (P6)

3.3 Masks

The vast majority of APs discussed the use of masks in the interventions. For example, one participant suggested that "the metaphor of the mask is central, pivotal to success, they really connect with people" (P5), with another coining it to be their "signature tool" (P2). It was explained that masks can represent individual identities; an AP would talk to groups about which masks each individual felt they identified with. Most APs commented on the effectiveness of fragment mask exercises, for example, one AP stated, "They worked amazingly, so simple but effective, people invested in challenging a character that they would identify with and then challenging them with it" (P1).

It was reported that fragment masks get the most positive feedback from groups attended by participants with an offending history. For example, one AP commented, "…fragment masks – they're always a success, with any group" (P4). Another AP commented that when using masks, they have felt a shift in people, "an aha moment" (P3). Similarly, another AP stated, "…when they do get it – it's like a penny drops, and it tends to stick with them" (P5).

One participant discussed their thoughts around the accessibility of the masks exercise believing that it is an easy concept for clients to grasp: "…very few have struggled with the concept of masks" (P5).

Participants went on to describe why they felt the masks were beneficial for clients. One AP articulately reported that the mask operates on metaphor and symbols, and that it taps into a language that connects with people visually and physically. Further comments included: "…masks move understanding and meaning past cognitive stuff to something deeper" (P2), and "Mask work allows people to shift, to externally gain an objective view of themselves, to stand outside oneself. A 3D version enabling self-analysis" (P7). Another AP commented that in the mask exercises, "People can resist the labelling and stigma; it encourages communication…" (P7). In short, it was reported that the masks exercises were a particularly powerful and accessible tool for GTC interventions.

Theme 4: Challenges and Support

4.1 Challenges

APs mentioned challenges around the emotional impact that the work had on them and the difficulties of working within a prison environment.

One AP spoke of their work with what they termed "lifers (P2)" and the difficulties in hearing about serious crimes. The AP described this as a "weird out of body experience, listening to graphic details of sexual offences; the crimes were shocking, so they were memorable" (P6). The AP went on to report that images they had when listening to this caused them to experience a type of cognitive dissonance, meaning there was a conflict between not wanting to work with people who had committed such offences, but on the other hand wanting to help people. The AP went on to report that this was very emotional for them:

"I had to work with sex offenders, which wasn't something I wanted to do" (P6). Another commented:

> I worked with young people, I started out passionate, but it ended up being heart-breaking listening to their stories and challenging to engage them in the short term as their lives were so chaotic and unstable. It was hard to get outcomes and keep the group together. (P5)

They further spoke of dealing with difficult group members, and clients in denial of their offending behaviours, and how tricky it is to challenge that. One AP mentioned that the volume of emotive stories they hear can be difficult to cope with, commenting "…the emotional side of the work – the relentlessness of the suffering. You get to know about peoples' lives, you meet lots of people so there is like an accumulation of all of those stories" (P2). Another commented that it was hard to forget the more negative stories and experiences clients had, saying that some of the more memorable experiences were the "dark ones" (P2). One AP stated that, ultimately, the work "…takes its toll" (P6).

An external challenge that was cited by APs was the challenge of working in the criminal justice system and the frustration this brings. For example, one AP commented, "sometimes the system can undo the great work we do – i.e., prison guards" (P3). The AP went on to describe work they had done with a young offender, who was then collected from the group by a prison guard who expressed negative views about the DTP to the young person. The AP commented that this had a negative impact on their (the APs) motivation for the job; feeling that their work may be undermined or undone by others.

Others mentioned certain practical challenges that come with delivering interventions in a prison setting, for example,"…you're fighting against the system…you can be in the middle of a project and there's an unpredictable lock-down, so you lose an afternoon session. You're not in control" (P3).

4.2 The benefits of Challenging Yourself

Participants commented on their positive experiences of being challenged in their roles, both for themselves and for the clients. It was commented that it made them better actors and increased their level of confidence in their abilities:

> When performing for people, you need to be authentic in it, I played someone with substance misuse problems. You've got to be at top of your game. This confidence came from the positive feedback I had received from both colleagues and clients, and it made me want to do it again and again. (P4)

Another AP suggested "the role was a life apprenticeship for me, it made me want to be the best actor I could be" (P2). In addition, they discussed the cyclical process of seeing the positive impact that their work had on individuals, which led them to feel the role was rewarding, which, in turn, increased their level of motivation for their role. For example, one AP stated:

> …to see them develop through the programme - it's a beautiful thing to observe and be a part of. It's changing peoples' lives, it's escapism for them, they take a risk and it's so rewarding…when you're out and about in public and you see someone you worked with and they're doing well, and they say – 'wow man that was brilliant'… seeing positive change keeps you motivated to do more good (P2).

Similarly, another AP commented, "the role was challenging but rewarding, and once I saw how people responded I felt committed to helping people" (P3).

4.3 Training

In terms of the training APs received within their role, APs reported how valuable they found the training process. They described a lengthy, detailed, rigorous, interactive, and

intensive 6-month probation training process. All APs commented on the training they received. For example, one AP mentioned “the company deeply invest in their recruits’ training and development. GTC are at the top of their game”; “it was the best training I’ve ever had - it was life changing”; “there was a real integrity to it”, and “I kept learning throughout the time I was with GTC” (P2).

Many APs discussed how the training had served them in their subsequent roles: “I still use and value the training I received today the work I do now is shaped from the training I had at GTC” (P6); “…[I’ve] never forgotten the training, it has served me brilliantly since” (P6); and “…unique and special training, transferable to other things” (P6). Another commented “My training was a long time ago, but I have clear memories” (P1).

APs commented on the good structure and format of the training. For example, “…it was comprehensive. There was a good theory practise balance to the training” (P5) and “the training provided me with a strong theoretical grounding in social learning and CBT” (P5). APs further commented that the training was not easy but that it was worth putting in the effort for: “…the training was very hard but rewarding, you need to retain a lot of information” (P7); “…it can be overwhelming, hard, and it’s very hands on” (P7); and “…it requires a lot of engagement” (P3). An AP discussed that the processing of taught information was new to them as they had to remember a lot and manage situations, commenting “it was a quick rehearsal period, so that was challenging, your brain learns the process after a while, the brain adapts”, and “You learn how to be facilitator - managing conflict and challenges” (P4).

One AP reported that there was no “magic wand” (P4) when it came to the training; they likened it to planting a seed that will grow later. The training is, in effect, ongoing in that it extends beyond the first six months, involving practical experience, support and reflection

(see section 4.4). They went on to discuss that it is a question of watching how the process works, then developing, for example, your own mask work:

> …you initially watch performances - group work, observation stuff, then slowly integrating you into the work. You then get small roles in performances, as you progress you get bigger roles, then fragment masks. As time progresses, you know you're doing ok. (P1)

4.4 Learning on the job

APs reflected that a lot of what they learnt and achieved was largely as a result of observing more experienced colleagues, being supported by them, and engaging in a process of reflection.

One AP stated, *"*80% of the job is observing, 20% theory…Initially I felt massively out of depth, it doesn't work for everyone, but performance was natural to me" (P2). Another stated:

> I loved it straight away, senior practitioner was so skilled with masks – they used their skills to demonstrate cognitive distortions versus speaking from the heart (mask up, mask down). Watching performances, group work observation stuff, then slowly integrating you into the work…if you're struggling – senior staff step in and help in building confidence. You start by having small roles in performances, as you progress you get bigger roles. You learn about the fragment masks, and theories. Challenging myself professionally, cool experiences would happen a lot. (P4)

APs spoke in depth about the high levels of support they received from the first day and the value of having a senior member of staff or mentor allocated to them. As the GTC is a touring theatre company, APs reported that they spent a lot of time with colleagues, which helped build bonds. APs reported on the level of supervision received and they all praised this

element suggesting, “The level of skill of senior staff can be intimidating as they are so good, but so inspiring” (P6). APs described the supervision element as ongoing and thorough.

They also spoke of the formal process of reflecting after practice/each piece of work and keeping a reflective diary. “Reflections can be exhausting and emotionally difficult, listening to people with mental health issues and why they’re inside , their energy, anger, pressure, anxiety, heavy emotions…” (P1). Furthermore, one AP commented:

> Being part of a tight knit group/company helps with the emotional side of the work – the relentlessness of the suffering. You get to know about peoples’ lives, you meet lots of people so there is like an accumulation of all of those stories. (P7)

All APs reported on the process of reflection, and the GTC’s commitment to learning from experience: “It’s rare in the arts to have such great support/supervision. You observe others a lot and receive structured feedback” (P3). APs reported on the structure of feeding back to trainees, suggesting that there is a self-report/verbal review to complete following all pieces of work, which encouraged reflection on: this is what you need to work on; what did you notice about others’ performances; what you felt went well; what would you do next time, as well as what the thoughts of others were, to promote taking feedback on board and to identify areas for development.

Discussion

This study aimed to explore the views and experiences of both former and current GTC Actor Practitioners (APs) who delivered DbTs to individuals in forensic settings. APs discussed, for example: aspects of the programmes they felt were successful; techniques used to engage clients; the therapeutic relationship; their own training; and challenges they have

experienced. Four key themes and a number of sub-themes were identified following Thematic Analysis of the transcripts. These themes are discussed below with reference to relevant literature.

Regarding the first theme, *The client journey*, AP responses mirrored, in part, academic literature regarding the beginning stages of the process of behavioural change (e.g., The Stages of Change, Prochaska & DiClemente, 1983; Multi-factor Readiness to Change Model, Ward et al., 2004; Readiness to Change Framework, Burrowes & Needs, 2009) in terms of their experience of barriers for clients engaging with interventions (e.g., negative attitudes, negative behaviours, and a lack of trust). Readiness to change literature (Ward et al., 2006, found in; Willis, Prescott, & Yates, 2013) suggests that having a positive view of the intervention and the therapist (i.e., being willing to trust them) is essential in order for an individual to feel ready to engage with an intervention. Some APs mentioned that, in attending the intervention, an individual was showing some recognition that they needed to change and a willingness to take action to change. Nonetheless, there was also a general consensus that many clients would be guarded and put up barriers. It was felt that clients were sometimes pessimistic about how the intervention may help them and displayed behaviours (e.g., putting on a tough front) which would make engagement with the exercises problematic. The pessimism and suspicion, in some cases, was felt by some APs to be as a result of a lack of trust in them and the intervention. Positively, however, and as alluded to in the results section, once those barriers were overcome, significant changes were observed within participants who continued to engage with the intervention.

Thompson (1999) noted that clients may have additional complex issues to overcome prior to being able/willing to engage with DbTs. The issue of a lack of trust in APs and the intervention could be viewed as a primary issue, which APs are aware of and need to address in order to increase the likelihood of engagement. Research has reported on the importance of

trust in therapeutic environments. For example, Papagiannaki and Shinebourne (2016) report on the importance of not only trusting the therapists but also trusting that the intervention provides them with a safe space where they are able to be themselves and open up. It is apparent that APs in this study were acutely aware of the range of potential barriers which clients may experience and the reasons underlying these barriers (i.e., past experiences, negative thought processes). Having this knowledge and understanding (as a result of training and experience) allows APs to support clients in the process of change through promoting engagement with the intervention. This suggestion is in line with the findings of Stephenson and Watson (2018) whereby clients reported that, as they got to know the APs, they felt able to trust them, which they felt helped them benefit from the intervention.

In addition to trusting the APs, it was noted that engagement with the intervention increased when clients felt they were in a safe space. Exercises led by APs put an emphasis on working collaboratively with one another; sharing and making meaning together. This supports previous research, which has highlighted that clients need to feel safe within a group in order to talk about their experiences (Casson, 2004; Papagiannaki & Shinebourne, 2016). Within the issue of having a safe space, there was a more nuanced discussion around the group dynamic. For example, it was noted that, when dominant members of the group were positive and engaged, this would have a positive knock-on effect for the rest of the group, who would follow their lead. This is concurrent with literature which has found that dominant members of the group can be the key that unlocks cooperation from the rest of the group; in other words, watching more dominant people engage may motivate quieter members of the group to take part (Keisari et al., 2018), boosting their self-esteem and confidence to do so (Cassidy et al., 2017). However, APs also noted that clients may feel uncomfortable if there are dominant group members who have a negative outlook. For example, clients could fear they may be laughed at or mocked within the group; an issue which has been linked to

feelings of inferiority and embarrassment, making it more likely that a client will withdraw from the group (Keisari et al., 2018).

APs expressed their awareness of the types of group dynamics, which occur in their interventions. Due to the flexible and responsive nature of the interventions, they are able to use their skills and training to address communication within the group and create a safe space where clients feel that they are respected by the APs and their peers; an issue which has also been highlighted in previous research as being key to engagement (Davey et al., 2014). Teasdale (1997) also reported that key to DbT within forensic populations is a supportive, community treatment ethos, which, if sustained throughout an intervention, can reduce the likelihood of future recidivism.

Essentially, these points refer to the role that improved communication skills can play in the process of change. APs reported that the communication skills participants acquire and practice in the group can be used to communicate more effectively with others in their life. Being in a group can create great bonds and connections with group members, which can further enforce feeling safe within a group and can help people recognise similarities between themselves and others (Baldwin et al., 2007). Clients can gain a sense of identity and community as well as an understanding of how they are viewed by others (Bourgon, 2013).

It is clear from the emphasis APs place on collaboration and interaction with other group members, that this is a key mechanism by which clients gain insight into their own behaviour and the behaviour and experiences of others. APs noted that clients could learn tolerance and respect for each other and be better able to empathise; skills which are transferable to situations outside of the group setting. Stephenson (2016) carried out an evaluation of two GTC interventions using a client sample of female offenders (N = 21). Client views captured in the Stephenson (2016) study, reflect what was reported by APs in the present study. For example, clients in Stephenson's (2016) study also reported that: they

felt better able to communicate with others; they had increased levels of self-esteem and confidence; they gained insights about themselves and others; they bonded with group members; and felt that they changed for the better.

Regarding the second theme, *APs qualities*, participants spoke about the circumstances that had led them to join the GTC. They discussed the qualities they had prior to the role as well as those they developed whilst they were working for the GTC. A key finding across the data was the extensive training provided by the GTC and the level of support APs received throughout the course of their employment. Many former APs reported that they continued using skills developed from this training in their work today. It was clear from responses that APs felt a strong affiliation and affection for the GTC; possibly because they were aware of and grateful for the time and investment GTC made in them individually. This could potentially link to the length of time they spend/spent with the GTC and the commitment they expressed to both the GTC and their role.

Many reported having studied for a degree in drama/theatre, commenting that they liked that they could use their drama education to help vulnerable people. The role brought together their interest and curiosity in criminal behaviour; their desire to work therapeutically within social change/justice settings, and their passion for theatre. One participant coining the term *therapy theatre.*

APs highlighted the need for clear *communication* in building a rapport with clients. A further key finding in this theme related to the skills and aptitude APs felt they needed when building rapport with clients. It is reported that sound relationships and positive alliances between the client and therapist result in more successful therapeutic outcomes (Baldwin et al., 2007; Casson, 2004; Kozar & Day, 2012). APs in this study echoed this, reporting that key to creating a safe space was building a strong therapeutic alliance with the clients. In line with the findings of this study, Lessner Listiakova et al. (2014) emphasise

using active listening skills to demonstrate to the client that the therapist is listening by using open body language and making eye contact.

Lowenkamp et al. (2012) looked at the therapeutic alliance between clients and their therapists in secure facilities when determining treatment outcomes and criminal recidivism. Lowenkamp et al., (2012) linked this back to practitioners being more aware of enhancing the quality of their process of assessment and intervention. They highlight the importance of a working case plan: continuously evaluating the case plan's aims and methods of implementation, as well as to be individualised. The assessment process can overemphasise risk assessment, and neglect responsivity. They reported that key to successful treatment outcomes were the interpersonal skills of the facilitator - being approachable and being firm but fair.

Key to this theme was also the concept of empathy for clients. Many APs highlighted that they believed some people in life lacked the opportunities afforded to others in society. This would suggest that APs felt empathy for people who have not had the same opportunities/advantages that they have benefitted from. Lessner Listiakova et al. (2014) proposed that therapist characteristics need to incorporate: empathy (i.e., putting themselves in their clients' shoes), emotional intelligence (i.e., knowing when to step back, and when to push forward in interactions with clients, as well as remembering to respect their knowledge and experience), emotional engagement, professional distance, and stability.

Bougoun (2016) reported that successful engagement in interventions was linked to the efforts from both the client and therapist in understanding each other's perspectives. These findings support the current study's findings whereby APs highlighted key skills they believed to be helpful in building rapport and trust in the therapeutic relationship. This included: to be authentic, down to earth, and to treat clients as equals. In doing so, this demonstrates a two-way respect to facilitate a therapeutic alliance and promote engagement

with DbT. APs reported that strong collaboration led to obtaining knowledge of a client, which was helpful in creating performances that were more authentic (i.e., that reflect the client's experiences, and lead to opportunities in challenging them). APs reported that some clients can be intimidating, or guarded, but by the APs putting in effort to find common ground, clients perhaps felt that the AP had a genuine interest in them. APs reported that ice breaker exercises and making efforts to talk to clients during break times helped with this too.

Lastly, meaningful therapeutic work is a key component in the therapeutic relationship (Bougoun, 2016). In line with findings from the present study, research has shown that actively showing clients that the therapist believes in and cares for their work can be hugely beneficial in promoting client engagement. Further, the APs in this study reported that showing confidence in their work and that they are not afraid to make mistakes can also improve the therapeutic alliance. These findings are in line with Listiakova et al. (2014), who developed the self-evaluation tool for drama therapists working with clients. This tool provides a reference point in terms of competencies linked to the therapeutic process. Interestingly, and in line with findings from this study. Listiakova et al. (2014) report that humour can help with rapport building and engagement. They go on to report that humour should be considered once trust has been established between the client and therapist.

Regarding the third theme, *creativity vs manualised interventions*, APs spoke of the flexibility and freedom they as APs have/had, as opposed to other professionals delivering more formalised, manualised interventions. DbTs delivered in prison settings arguably have some similarity with manualised OBP programmes, that is, both are grounded in the CBT approach, the Risk-Needs-Responsivity Model (RNR) (i.e., addressing the needs of clients and tailoring the delivery to the individual), as well as "What Works" literature (i.e., best practice for working with offenders). However, it is perhaps apparent that DbTs (e.g., applied theatre) are better placed to address psychological barriers to change (through the use of

masks, for example) and to be responsive to the sometimes diverse needs of individuals. As highlighted in previous research, there is a lack of creativity in manualised interventions; an issue which may lead to lower engagement with the programme (Davey et al., 2014).

APs discussed that DbT can be responsive and adapt during the course of the intervention based on the needs of the group and/or individual participants. Should someone discuss a personal situation, the APs can ask the client what that situation looks like and then act it out for the group. Clients can come to a realisation or make a decision which is specific to their own pathway; making things feel more relatable for an individual. They highlighted that GTC DbT is less conventional than manualised programmes and is un-school-like, making it less formal with no traditional teaching element. They went on to report that DbT caters for different learning styles, and that DbT can be more effective for visual and active learners, and for people who have low attention spans; something which has also been highlighted in previous literature (Smeijsters & Cleven, 2006). In short, APs reported that core to the perceived efficacy of GTC programmes, is their responsiveness to the needs of the client, and that the ability to be flexible and tailor interventions to the needs of the individual is distinctive to arts based therapies.

This theme further explored APs views on the impact their creative performances have on clients, reporting that clients feel entertained and inspired when watching the performances. More specifically, they highlighted that the clients were able to make personal connections and find a different or new identity (i.e., not just thinking of themselves as a prisoner). They highlighted that DbT is a unique platform for clients to explore thoughts, feelings and behaviours in a creative way. These findings are similar to that of a study exploring the views of a group of female prisoners with reference to another DbT programme (Reason, 2019). APs further reported that the arts can reach a part of a person that no other intervention can, connecting with a deeper part of themselves. APs emphasised the way that

GTC interventions help clients to have their own realisations about their lives rather than be didactic in their approach. This is in line with a quote from a 17th century philosopher taken from the GTC handbook: "People are generally better persuaded by the reasons which they have themselves discovered than by those which have come into the mind of others" (as cited by Baim et al., 2002, p.23), as well as more recent literature on DbT in general (Smeijsters & Cleven, 2006). APs discussed the use of masks in DbT, and the uniqueness of this practice, compared to more formalised interventions. One AP stated that masks move understanding and meaning past a cognitive level to something deeper. They explained that masks are the GTC's signature tool. APs discussed that the intervention taps into a language that connects with people visually and physically. Many APs mentioned fragment masks as an easy concept for clients to grasp; believing that very few people struggle with the concept of masks. APs explained that through mask work people can act out a character they identify with to gain an external/objective view of themselves, as also noted by Harkins et al. (2011). They explained that clients are able to create a fictional character, where they are able to view their behaviour and reflect from a safe distance (figuratively speaking), and that it is in this safe space where clients are able to reflect and learn about themselves and their behaviour. This finding is supported in more general literature regarding role-playing exercises in drama therapy, which allows clients to play out real-life scenarios where they can change their role giving them insight into others' perceptions of their behaviour. Chandler (1973) suggested that this was key in reducing reoffending rates in their sample of male delinquents (N = 45).

In addition to commenting on how GTC interventions relate to clients, APs also discussed their work in relation to the challenges that their role involved and the support and training they received in order to be effective in their role (see theme 4). As noted in previous literature relating to the consequences of working in a therapeutic manner with offenders in a prison environment (e.g., Friedtich & Lieper, 2006; Volker & Galbraith, 2018),

APs commented that, at times, they found it challenging to hear about crimes that clients had committed and the distressing details of their life prior to their arrest. APs described a dichotomy whereby their role was to help people who had, in some instances, committed crimes which appalled them (e.g., sexual offences).

Although not mentioned by APs in the current study, previous research has reported that therapists working with prisoners, who had committed sexual offences, felt anger, hostility and distance towards such prisoners (Friedrich & Lieper, 2006). However, in reference to challenges in the current study, there was more emphasis on the challenge of hearing about the suffering faced by clients; an indicator of high levels of empathy for clients (see theme 2). APs in this study reported finding it hard to forget the more negative, dark stories. Research suggests that therapists can encounter burnout and distress (Farrant, 2012); issues which were touched on by some APs in this study. Farrant (2012) also noted that listening to harrowing stories can lead to low job satisfaction. However, APs spoke of how rewarding they found the work to be due to seeing positive changes in clients. It was apparent that APs attempted to mentally distance the client from their offence, instead focusing on their client's life story and journey; a technique also used by other therapists working with this client group (Farrant, 2012).

APs also highlighted the challenge of working with clients in denial of their crime which, in turn, could mean the client is not taking responsibility for their actions. The challenge of rehabilitation of offenders who are in denial of their offence is well documented (Lord & Willmot, 2007). However, in the UK, being in denial of an offence does not preclude an offender from taking part in treatment programmes and, in some instances, offenders may come out of the denial stage (Lord & Willmot, 2007). The emotional challenges faced by APs in this study were evidentially mitigated by being part of a tight knit company that offer comprehensive training and ongoing support (see below).

APs highlighted that working within the constraints of the criminal justice system can lead to frustration. They stated that sometimes the system can undo the changes that have occurred in clients as a result of the intervention. One AP stated that it feels like fighting against the system at times, noting that there can be unpredictable lock-downs (within the prison setting), which can result in a session being lost. Some also felt a lack of control, at times, as a result of working in a prison. This can affect their motivation for the job; feeling that their work may be undermined or undone by others. Such challenges were noted by Bergman (the founder of the GTC) when their work began in 1980. Bergman (2017) reflected that despite the challenges of working inside the criminal justice system, it was not possible to elicit such change from the outside: "And so we went inside the systems…We made change in places where none had existed for decades. Unable to change the United States or the Unites Kingdom from the outside, we poured our energy inside these systems instead" (p. 324).

APs also reported on the benefits of the challenges of their roles. Some APs noted that their work made them better actors and increased their level of confidence in their abilities. They reported that this confidence came from the positive feedback they received from co-workers and clients. One AP stated that once they saw how people responded, it made them feel more committed to helping people.

APs reported that a large part of the role of an AP was observing more experienced colleagues on the job (i.e., learning on the job). Many APs reported that initially they felt out of their depth, but that training, support, and observations helped them feel more at ease with their role. For example, they reported that senior practitioners were very skilled with the use of masks (e.g., using their skills to demonstrate cognitive distortions versus speaking from the heart, referred to *as mask up - mask down*). Further disclosures included APs feeling intimidated but inspired by senior practitioners. APs explained that at the start of their role

they observed more experienced colleagues (e.g., watching performances, observing group work), and then slowly began to work in the role of AP. APs reflected on the process of learning about the fragment masks and theories. They discussed that reflection was a large part of the process, reflecting after practice/each piece of work and keeping a reflective diary. APs reported on the level of supervision they were given and how this was ongoing, with one AP commenting how rare it is to receive such great support in the arts. Supervision involved giving feedback on what they (APs) noticed about others' performances (e.g., what they felt went well, what they would do differently next time and areas for development). The benefits of supervision and reflection when working in a therapeutic role are well recognised despite the fact that such reflections can be exhausting and emotionally challenging (Volker & Galbraith, 2018).

Methodological Strengths and Limitations

Despite the initial research plan emphasising the need to have similar numbers of former and current APs, there were more former APs than current APs in the current study. This was due to the Covid-19 pandemic which resulted in current GTC APs being furloughed and therefore unavailable to be interviewed. A limitation of mainly relying on interview data from former APs was that, as some of the APs had not worked for GTC for a number of months/years, their memory of the role and experiences may have been less accurate. In addition, their reflections could have been influenced by subsequent practitioner roles. It is recommended that current APs be interviewed at a later date to add to the findings of this study. Although the final sample size was small, it was considered sufficient for a qualitative study using Thematic Analysis, and the small sample size allowed for an in-depth exploration of APs views and experiences. As outlined above and in Chapter 1 of this thesis, there are a range of DbTs available, and they do not follow a prescribed template per se. As such, the findings of this study cannot be generalised across all types of DbT. However, it is evident

that some findings are concurrent with previous literature which could offer support to some generalisations being made.

Furthermore, when deriving themes from data there is the potential for researcher bias relating to the interpretation of data. In order to somewhat mitigate this potential, the steps outlined by Braun and Clarke (2013; 2019) were closely adhered to; a reflective diary was kept in order to recognise biases, and themes/sub-themes were discussed with an academic supervisor as well as another doctoral student (for the purpose of assessing inter-rater reliability).

Implications for Practice

It is hoped that the results of this study will be of interest to other practitioners who deliver DbT in terms of providing them with, for example, information regarding potential barriers for engagement, helpful practitioner qualities and appropriate levels of support. More specifically, based on the findings of this study, it is suggested that the following issues be considered by both therapists using the arts and other professionals who facilitate group work with prisoners:

- Breaking down barriers to engagement: APs suggested this can be achieved through ice breaker exercises, communication (see below for detail) and, notably for the GTC, the mask technique.
- Therapeutic relationship: Despite the often short duration of interventions (e.g., one week), APs highlighted the need to build a good therapeutic relationship with clients. More specifically, they suggest finding common ground; being down to earth; not coming across as patronising; recognising that the client is the expert in their own life/experiences; displaying empathy; seeing past a person's crime and displaying passion for their work.

- Awareness of group dynamics: The impact that group members have on each other needs to be considered. Efforts should be made to help clients get to know and trust each other and build bonds.
- Practitioner qualities: Due to the challenging nature of the role it is suggested that certain qualities are important for practitioners to possess (e.g., empathy, passion, and drama/theatre skills).
- Creativity: Where possible, practitioners should be mindful that a *one-size-fits-all* approach should be avoided; interventions should be responsive to the needs of individuals. Key to this is gaining knowledge of a client's past experiences. Practitioners should consider the different learning styles of clients and use creative techniques to help clients come to realisations for themselves rather than an over reliance on didactic methods.
- AP support: Comprehensive training and adequate support should be provided for practitioners. This would include observing sessions prior to leading sessions, taking time to reflect, and forming bonds with colleagues.

Recommendations for Future Research

As previously mentioned, it would be beneficial to explore the views and experiences of APs using a larger sample size and with a higher percentage of current APs. This may be achievable in the near future. The analysis of additional data would further add to the knowledge base of the therapeutic values of GTC DbTs. In addition, it would be beneficial for there to be longitudinal studies on the impact of GTC DbTs to ascertain whether the interventions have a long lasting positive impact. Furthermore, it is suggested that further studies be conducted which explore both Drama based therapists' and participants' experiences (in terms of the therapeutic process) of a range of drama based interventions, in

order to provide further insight and evidence for the usefulness of drama and theatre based interventions within forensic populations.

Conclusions

This study has reported on the views and experiences of Actor Practitioners who work, or have worked, with the Geese Theatre Company in the UK. The findings have highlighted a number of issues relating to prisoners who take part in GTC interventions (i.e., barriers to engagement, and the benefits of the intervention). In addition, the study provides an overview of the qualities APs felt necessary for their role, APs' views on the differences between GTC interventions and manualised interventions (i.e., noting the advantages that creativity/flexibly affords), as well as providing an insight into the challenges they experience/d and the benefits of training and support.

Although this study could be considered to be nuanced in that it reports on the views and experiences of APs from a specific company, it is hoped that the findings will add to the knowledge base around DbTs more broadly. In addition, it is hoped that the study will shed light on factors which have resulted in positive outcomes and feedback from clients who take part in GTC interventions. Finally, it is hoped that this study would provide management level prison staff (i.e., those who make decisions regarding whether a GTC intervention would be beneficial for prisoners under their care) with an insight into the type of skills, knowledge and experience possessed by GTC APs.

CHAPTER 4

A critique of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS; Stewart-Brown et al., 2007)

Introduction

There are a variety of measures available which focus on the construct of *well-being*. Wellbeing is considered to be conceptualised through: a) affective-emotional; b) cognitive-evaluative; and c) psychological functioning (Tennant et al.,2006).

Research into mental well-being is a relatively recent area of study. As such, there remains controversy around the definition of mental well-being (Clarke et al., 2011) due to it being a broad term and covering a range of issues. In 2004, the Royal Society (as cited in Clarke et al., 2011), defined well-being as a "positive and sustainable mental state that allows individuals, groups and nations to thrive and flourish" (p. 2). Despite the difficulty in defining well-being, researchers have attempted to conceptualise it in order to help measure it for practice and research purposes (Mauss & Robinson, 2009).

Measures of psychological well-being

Existing measures of well-being include the Scales of Psychological Well-Being (SPWB; Ryff & Keyes, 1995), which measures hopelessness, purpose in life, life satisfaction and self-esteem. This tool focuses on eudaimonic[4] well-being and measures psychological functioning[5]. Furthermore, the Global Life Satisfaction Scale (GLS) aims to measure quality of life**.** The Euroqol Health Status Visual Analogue Scale (EQ-5D VAS; Brooks et al., 2003) aims to measure health states using a visual analogue scale. However, the SPWB, GLS and the EQ-5D VAS do not specifically measure affective-emotional[6] and cognitive-evaluative[7]

[4] Eudaimonic well-being refers to living a life of purpose in striving for human completeness, this includes self-actualization, and individual expression (Carruthers & Hood, 2004).

[5] Psychological Functioning or health refers to the internal functioning of a person's emotions, behaviours, and social skills which encompasses their mental health as a whole (Storbeck & Clore, 2007).

[6] Affective Cognition, the merging of emotional and cognitive processes to produce a behaviour (Mauss & Robinson, 2009).

[7] Cognitive Evaluation is a term used to explain the impact external consequences can have on internal motivation (Storbeck & Clore, 2007).

functioning**.** This is a noteworthy criticism because emotional processing as well as cognitive reasoning are important when considering a comprehensive concept of well-being. In comparison to the aforementioned measures, the Satisfaction with Life Scale (SWLS; Diener et al., 1985) sets out to measure cognitive-evaluative facets of mental well-being such as "I am satisfied with my life". In addition, the Personal Well-being Index (PWI; Tomyn et al., 2011) aims to measure subjective well-being, tapping into an individual's satisfaction of one's self. In summary, whilst the aforementioned instruments measure the construct of mental well-being, they use slightly different conceptualisations and/or look at isolated aspects of well-being. Tennant et al. (2006) highlight limitations of these measures and call the reliability and validity of the measures into question. Of particular note is that the Personal Well-being Index (PWI) has mixed reviews despite the fact that it has been used with a number of populations globally, as not all of the constructs within the tool are thought to be understood across different cultures (Lau et al., 2005). For example, the understanding of happiness is subjectively different in non-European cultural groups compared to European cultural groups (Taggart et al. (2013).

Rationale for critiquing the WEMWBS

The WEMWBS – designed as an alternative to the above measures - was developed to measure facets considered important to the construct of mental-wellbeing**.** The WEMWBS is now widely used by health care professionals and governance (policy development) within the UK (Tennant et al., 2006, 2007) and has been validated and used in many other countries (i.e., Italy, Spain, Germany, France, Netherlands, Belgium, Iceland, India, Pakistan, Malaysia, and South Africa) (Gremigni & Stewart-Brown, 2011; Ikink et al., 2012; Lloyd & Devine, 2012; Lopez et al., 2013). It has been used to assess mental well-being across a range of populations such as school students (Clarke et al., 2007), minority ethnic groups (Stewart-

Brown et al., 2011), and prisoners (Stephenson & Watson, 2018). Due to the wide usage of the WEMWBS, there is a need to ensure that it is a reliable and valid measure to allow practitioners to make accurate assessments of an individual's well-being. Existing research findings on the WEMWBS have not been synthesised to date. The following review aims to describe the WEMWBS's theoretical underpinnings, development, administration, and scoring. The review will also explore the psychometric properties of the WEMWBS using guidelines set out by Kline (1986), focusing on the validity and reliability of the tool. This review will draw from current research to inform a discussion of the psychometric rigour of the WEMWBS.

Overview of the WEMWBS

Tennant et al. (2007) set out to create a robust measure of mental well-being (WEMWBS) and to validate the measure across a range of populations. The WEMWBS was developed to measure facets of mental well-being: hedonic[8] (i.e., subjective happiness, life satisfaction) and eudaimonic (i.e., psychological functioning, self-realisation) (Tennant et al., 2007). The WEMWBS is a self-reported measure which consists of 14 questions that connect to an individual's state of mental well-being (i.e., thoughts and feelings) across the previous two weeks. The measure uses a 5-point response scale (ranging from *none of the time* to *all of the time)*. Questions are worded positively and were developed to map onto attributes of mental well-being (hedonic and eudaimonic). Item statements include, for example, "I've been feeling interested in other people"; "I've been feeling relaxed", and "I've been feeling optimistic about the future" (Tennant et al., 2007, p.7). The measure does not contain questions regarding spirituality or purpose in life, as these were considered by the authors to

[8] Hedonic wellbeing refers to the notion that happiness is dependent on increased pleasure and decreased pain (Carruthers & Hood, 2004).

be less related to mental well-being (Tennant et al., 2006, 2007). The measure sets out to assess mental well-being as opposed to individual factors that potentially determine mental well-being (e.g., relationships, dealing with conflict, and problem-solving) (Maheswaran et al., 2012). In addition, the authors further highlight that the tool does not aim to measure well-being as a function of socioeconomic factors (e.g., unemployment and low income group). The authors state that key components of the measure include: the emphasis it places on its normal distribution within the general population (with no floor[9] or ceiling effects[10]); and its applicability to healthcare professionals working clinically and informing policy at a governance level (Tennant et al., 2007). WEMWBS results are reported as a mean score for a specific population, or they can be shown as a range of scores within a given sample. In addition, a calculation of the variance of the scores within the sample (either whole or sub-groups samples) is reported using 95% confidence intervals, or the standard deviation.

Development of the WEMWBS

The WEMWBS was developed based on the Affectometer 2 tool (a measure developed in New Zealand in the 1980's (Kammann & Flett, 1983)) as well as data from systematic literature reviews on the topic of mental well-being and views of a multidisciplinary expert panel. The developers of the WEMWBS set out to validate the Affectometer 2 in the UK, which formed the initial stage of their development of the WEMWBS. The rationale for developing the WEMWBS from the Affectometer 2 was due to some shortcomings of the Affectometer 2, which were identified in the validation study and

[9] Ceiling effects – a ceiling effect occurs when a person scores the highest or near highest score on all items of a measurement tool. This does not indicate whether the person's true level of functioning has been accurately measured. (Lim et al., 2015).

[10] Floor effects – when a large proportion of the sample scores the minimum value on as measure. This skews the distribution of scores, which impacts on the ability to differentiate between clusters of individuals (Lim et al., 2015).

are outlined below. It was reported that high levels of internal consistency ($r = 0.94$)[11] [12] suggests a reluctance to social desirability bias[13]. One of the shortcoming mentioned was and length of the Affectometer 2 (40 *items* – 20 positive, 20 negative) - in items or the time taken to complete it. A further purpose was to develop a mental well-being tool with a single underlying construct, encompassing areas of well-being (e.g., affective-emotional, cognitive-evaluative, and psychological functioning).

The Affectometer 2 (as with the WEMWBS) comprises of hedonic and eudaimonic facets of well-being and has been shown to correlate to mental health, balancing both positively and negatively worded questions. This measure was reported to be used by health practitioners/researchers based in the UK (Tennant et al., 2006; 2007). For the Affectometer 2 scale development, nine groups of eight adult participants in England and Scotland (N = 56) were recruited. Participants were identified from a range of community groups covering various demographics (i.e., gender and socio-economic status). One of the groups consisted of mental health service-users (Tennant et al., 2007). From this initial validation study, and as reported above, the Affectometer 2 measure was reported to be lengthy. Both the Affectometer 2 and the WEMWBS are self-reported measures which can be susceptible to social desirability bias. Investigation into internal consistency of the measure proved to be high ($r = 0.94$), meaning that the questions in the measure were measuring the overall shared construct of mental well-being (Stewart-Brown et al., 2009). The Affectometer 2 was shown

[11] A good coefficient – refers to a range of values that either show a positive or negative correlation. A correlation of -1.0 displays an absolute negative correlation, and a correlation of 1.0 displays an absolute positive correlation. The values range from -1.0 and 1.0. A value larger than 1.0 or below -1.0 shows an error in the correlation measurement (Ratner, 2009).

[12] High/good internal consistency – refers to measures where several items that propose to measure the same general construct produce similar scores (Revicki, 2014).

[13] Response bias – refers to the tendency for people to respond incorrectly or inaccurately to measure items (Wetzel et al., 2016).

to have: high face and construct validity; good test-retest reliability over time; and appropriate differential status between population groups (Tennant et al., 2006, 2007).

A focus group study was further conducted with participants from a variety of socio-economic backgrounds, which showed that the measure had a good balance of both positively and negatively worded questions. However, some of the questions were considered difficult to understand, and some participants considered it to be a measure of mental illness rather than of mental well-being (Tennant et al., 2006). Based on these findings, Stewart-Brown et al. (2007) devised a number of questions for their own measure - the WEMWBS. Some questions were taken from the Affectometer 2 (excluding the difficult and negative questions identified). Further, they reduced the number of questions in line with the positive mental well-being literature (i.e., affective-emotional aspects, cognitive-evaluative facets, and psychological functioning). This resulted in a 14-item scale, namely The Warwick-Edinburgh Mental Well-being Scale (WEMWBS).

Administration and scoring

The WEMWBS is a self-report measure and participants completing the WEMWBS are asked to choose a score ranging from 1-5 (*1 = none of the time*, *5 = all of the time*) that most accurately illustrated their experience of each of the 14 items (see Table 5), based on the last two weeks.

Table 5

The Warwick-Edinburgh Mental Well-being Scale (WEMWBS; Stewart-Brown et al., 2007)

	WEMWBS Questions
1.	I've been feeling optimistic about the future

2. I've been feeling useful

3. I've been feeling relaxed

4. I've been feeling interested in other people

5. I've had energy to spare

6. I've been dealing with problems well

7. I've been thinking clearly

8. I've been feeling good about myself

9. I've been feeling close to other people

10. I've been feeling confident

11. I've been able to make up my own mind about things

12. I've been feeling loved

13. I've been interested in new things

14. I've been feeling cheerful

A total score is calculated by summing scores for all 14 questions. Therefore, the minimum achievable total score is 14 and the maximum is 70, with equal weights. A higher WEMWBS score suggests a higher functioning of mental well-being (Tennant et al., 2007). Due to difficulties in providing a concrete definition of mental well-being, there are no specific clinical cut-off limits on the WEMWBS, however Tennant et al. (2007) suggest that a three to eight point increase would suggest an improvement in mental well-being.

Scale validation: participants and data collection

Tennant et al. (2007) collected data from participants across seven disciplines from both undergraduate and postgraduate courses at Warwick and Edinburgh Universities (N = 2429). Participants gave information on their demographics such as: subject of study, gender and age. Participants completed the WEMWBS, along with two to four other randomly

selected scales from a possible eight other scales considered to be related to the construct of the WEMWBS. These included: The Positive and Negative Affect Schedule (PANAS; Watson et al., 1988); The Short Depression-Happiness Scale (SDHS; Joseph et al., 2004); WHO (Five) Well-Being Index (WHO-5; Bech, 2004); Scales of Psychological Well-Being (SPWB; Ryff, & Keyes, 1995); The Satisfaction with Life Scale (SWLS; Diener et al., 1985); Global Life Satisfaction scale (GLS, Eurobarometer Life Satisfaction [http://ec.europa.eu/public_opinion]); or the Emotional Intelligence Scale (EIS; Malouff et al., 1998). The WEWMBS was randomly administered prior to or following completion of the other measures. The eight measures were used to inform the WEWMBS by comparing scores. The measures were selected as they were considered to cover similar dimensions to the WEMWBS or to contain elements linked to mental well-being, for example, emotional intelligence and general health. In addition to the measures of well-being, the Balanced Inventory of Desirable Response (BIDR; Paulhus & Reid, 1991) was further administered to detect response bias (e.g., overly positive responses; self-deceit).

To validate the data obtained from the aforementioned student sample (N = 354), a further dataset was obtained from a Scottish population sample (N = 2075) (Tennant et al., 2007) to allow for comparison. Two measures were administered: the Scottish Health Education Population Survey (HEPS; Health Education Population Survey, 2006), and the Well? What do you think? Survey (Braunholtz et al., 2006). These scales cover factors such as gender, age, views on self-health, employment, total income per household, and age of leaving education. The population sample were further assessed on mental ill-health using the General Health Questionnaire (GHQ-12; Goldberg & Williams, 1988) which explores general levels of happiness, depressive and anxiety symptoms, and sleep disturbance, based on the past four weeks. The findings of these validation studies will be considered below.

Psychometric properties of the WEMWBS

Level of measurement

As mentioned above, the WEMWBS uses a numerical Likert scale format, which is considered preferable as a measurement tool, as opposed to only *yes* or *no* responses (Maheswaran et al., 2012). A benefit of the 1-5 Likert scale is that it has a mid-point (e.g., *3 = some of the time*), which allows respondents to provide a neutral response. However, it is important to consider that mid-points can produce false high ratings (Kline, 2000). Kline (2000) explained that by focusing on items relevant to the construct being measured there should be less use of an undecided or neutral option to questions.

Reliability

Researchers argue that both reliability and validity are equally important for a measure to be deemed valid (Maheswaran et al., 2012). For psychometric measures, it is widely acknowledged that reliability has two strains: internal consistency, and test-retest reliability (Kline, 1993). These two types of reliability will be discussed below with reference to the WEMWBS. Reference will also be made to the shortened (7 item) version of the WEMWBS for information only.

Internal consistency

Internal consistency is concerned with the correlation between items on a scale/measure; in other words, whether different items of a scale measure the same construct, and in this case, mental well-being. A sample size of less than 200 individuals is required to test the internal consistency of a measure (Maheswaran et al., 2012). Cronbach's Alpha (α) is commonly used to measure internal consistency (Klein, 1998; 2000). It is understood that for scores on a measure to have acceptable internal consistency, Cronbach's Alpha (α) must range from .60 to .70 (Klein 1998; 2000). High internal consistency is considered beneficial

as it suggests that the items within a tool all measure the same construct. However, if Cronbach's Alpha is too high then this may indicate that there are redundant items (i.e., repetition of items) or could indicate questions are too similar (Klein, 1998; 2000). Boyle (1991) stated that "…moderate to low item homogeneity is actually preferred if one is to ensure a broad coverage of the particular constructs being measured" (p. 6). If a scale only has two items for example, it will have a low (α) value, suggesting that (α) is a function of the number of items on a scale (Taber, 2018). Tennant et al. (2007) reported Cronbach's Alpha for scores in the aforementioned student sample (α = .89) and Scottish sample (α = .91), much higher than the previously suggested acceptable level of α = .60 to .70 (Klein 1998; 2000).

Stewart-Brown et al. (2009) conducted a Rasch analysis of the WEMWBS with a Scottish sample from the Scottish Health Education Population Survey (N = 779). The Rasch model looks at whether a scale measures the same way across any sample being assessed (Rasch, 1977). In the first instance, they reported a poor fit to the model. As such, a few items, which showed gender/age biases were deleted (e.g., "I've been feeling good about myself" and "I've been feeling cheerful") and led to the development of a seven item version, the Short Warwick Edinburgh Mental Well-Being Scale (SWEMWB; Stewart-Brown et al., 2009). The 14 item WEMWBS showed very high internal consistency which could indicate questions are too similar (α = .95). Although, the shorter version showed to have acceptable internal consistency, concerns were noted that reducing items on the tool may lead to a less robust measure of mental well-being (Tennant et al., 2007). Stewart-Brown et al. (2009) reported the SWEMWBS to be mainly free of item bias. Similar to the full length measure of the WEMWBS, higher scores within items indicate higher overall mental well-being.

Gremigni and Stewart-Brown (2011) tested the internal consistency of the 12-item Italian version of the WEMWBS using Cronbach's Alpha with both males and females, and

all ages of people. They used Intra Class correlation Coefficient (ICC), an alternative to Cronbach's Alpha. Gremigni and Stewart-Brown (2011) found no significant differences between the two groups (both from Italy, one population sample recruited based on number of years in education N = 345, and the other group recruited from a GP surgery N = 52). In terms of the internal consistency of the 12 item Italian version, the first group showed a Cronbach's alpha value of α = 0.86, and the second of α = 0.83, demonstrating good internal consistency.

López et al. (2012) adapted the WEMWBS into Spanish and investigated the metric properties of the tool. A student sample (N = 148) completed the WEMWBS along with similar scales measuring positive aspects (the same as the British ones). Internal consistency was investigated and was found to be high (α = .92). Of the 148 participants, 52 completed a one week retest. Four items were adjusted to be more transferable in terms of language related to the original measure. Overall internal consistency of α = .90, item-total score correlations (α = .44 to .76), and test-retest Intra Class Correlation Coefficient (ICC) (.84) were deemed adequate. Exploratory Factor Analysis (EFA) revealed some highly correlated dimensions, which led the authors to conclude that the Spanish version was more one-dimensional. one explanation for this was that 4 items had to revised to be as close to an equivalent to the British version as possible.

Positive scales (i.e., WHO-5; Bech, 2004, and PANAS-PS; Watson et al., 1988) showed high positive correlations with the Spanish version of the WEMWBS. Whereas the correlation with the Satisfaction with Life Scale (SWLS; Diener et al., 1985) was lower than shown in the original UK validation study. Similarly, Castellví et al. (2013) investigated the internal consistency of the WEMWBS with a Spanish sample taken from the general population (N = 1900) and high internal consistency was found (α = .93).

Furthermore, Taggart et al. (2013) investigated the cross-cultural rigour of the WEMWBS in measuring mental well-being among two UK minority groups (English speaking Chinese and Pakistani individuals). They looked at how well the WEMWBS withstood being converted into a different language, in addition to the Spanish and Italian versions mentioned above and wanted to investigate how the scale was viewed across different cultural groups. Once the WEMWBS was translated, it was administered alongside other measures designed to measure similar constructs (i.e., The Positive and Negative Affect Schedule PANAS; Watson et al., 1988). They used a method called *Rasch Modelling* (Rasch, 1977), created to assess psychometric properties of scales. Data from both samples showed an inadequate fit of the WEMWBS properties to the assumptions detailed in the model. The assumption was that the psychometric properties of the WEMWBS would be replicated when applied to minority ethnic English speaking groups (who identified as UK citizens) when compared to the wider UK population.

Consequently, the tool was adapted by items being introduced/removed to obtain a better fit (Kéry & Royle, 2016), using Exploratory Factor Analysis. One factor (e.g., "being interested in other people") was significant and explained 52% of the variance for the Chinese sample, whereas three factors (e.g., "I've been feeling optimistic about the future"; I've been feeling confident") were significant for the Pakistani sample (explaining 48% of the model variance). The authors concluded that the original version of the WEMWBS is representative of aspects of mental well-being relevant for the two ethnic minority groups they investigated.

Test-Retest Reliability

Test-retest reliability refers to whether the same scores are obtained when a test is administered on two separate occasions under identical conditions (Kline, 1998). Kline

(2000) purports that the time between tests be no longer than three months, and that the lowest correlation value of a good test is α = .80. Other theorists advise correlation values greater than α = .70 are sufficient to indicate good test-retest reliability (Streiner & Norman, 2008).

Test-retest reliability was assessed in the validation study by Tennant et al. (2007). A sub-sample of the student sample (n = unspecified), who previously completed the measure, were randomly selected to complete the WEMWBS again, after a gap of one week. Test-retest reliability for the WEMWBS after one week was deemed to be very good (α = .83). However, the authors suggested that test-retest be conducted to test the WEMWBS for stability of the measure over a longer time period than one week (Tennant et al., 2007).

Gremigni and Stewart-Brown (2011) measured the stability of the WEMWBS for Italian populations. Two groups from Italy participated, a general public sample (N = 345) and a second sample taken from a doctor's surgery (N = 345). The second smaller group from an independent sample was used for the purposes of assessing test-retest reliability of the WEMWBS (*n* = 52). The second test was administered one week after the first test had been completed. They reported good reliability characteristics for the WEMWBS over time (no α value reported). Similarly, López et al. (2012) conducted a test re-test study using an Intraclass Correlation Coefficient (ICC) with a Spanish sample (N = 52) and findings showed high reliability (α = .84). These findings were mirrored in Tennant et al.'s (2007) study, where they conducted a test-retest on a sub-group of their Scottish sample dataset at a one week follow up time point which showed that scores on the measure remained consistent.

Validity

For a scale to be deemed valid, in other words, whether it captures the concept it claims to measure, Kline (2000) purports that the following measures of validity should be

met: criterion validity, construct validity, content validity, and face validity. These are discussed below.

Criterion validity

Criterion validity refers to how well the outcome of one measure correlates to the outcome for another measure, both assessing the same underlying construct (Kline, 2000).

Tennant et al. (2007) used a Spearman's Rank correlation coefficient (r) test[14] to look at whether WEMWBS scores correlate with other scales measuring similar constructs (i.e., to ensure that similar tools of well-being are measuring independent constructs). They reported high correlations with measures pertaining to positive affect or well-being (e.g., SDHS (r = .73, p < .01); Joseph et al., 2004), WHO-5 (r = .77, p < .01); Bech, 2004), PANAS-PA (r = .71, p < .01; Watson et al., 1988), and SPWB (r = .74, p < .01); Ryff & Keyes, 1995). Further, moderate correlations were shown with the EQ5D-VAS (r = .43) (Brooks et al., 2003) which pertains to constructs of eudaimonic well-being and psychological functioning. In addition, significant correlations were shown between the WEWMBS and the two life satisfaction measures: Satisfaction with Life Scale (SWLS; Diener et al., 1985) and Global Life Satisfaction (GLS; Diener et al., 1985). Low to moderate correlations were shown between the Emotional Intelligence Scale (EIS; Malouff et al., 1998) and the WEMWBS (r = .48, p < .01) indicating that the two measures are not measuring the same underlying constructs. A moderate negative correlation was shown between the PANAS-NA (Watson et al., 1988) and the WEMWBS (r = -0.54, p < 0.01) suggesting that these two measures measure different constructs. In the population sample, the WEMWBS scores showed negative correlations with the General Health Questionnaire (GHQ-12; Rutter, 1997) (Ikink

[14] The r value is a correlation coefficient, it shows the strength of a relationship between variables.

et al., 2012; Tennant et al., 2007) which suggests that the WEMWBS is a standalone measure of well-being. These are expected outcomes that show that the measures are measuring similar constructs of mental well-being but that they're not too similar. Furthermore, Gremigni and Stewart-Brown (2011) conducted a two bi-factorial model[15] (independent and correlated factors) to ascertain how the WEMWBS correlated with other measures measuring similar hedonic and eudaimonic well-being constructs (e.g., PANAS; Watson et al., 1988, SWLS; Diener, et al., 1985, WHO-5; Bech, 2004, PGWBI; Dupuy, 1984). The independent factor was made up of items linked to the hedonic construct. The correlated factor was made up of items linked to the eudaimonic construct. Findings showed neither had a good fit to the data, meaning that the WEMWBS would appear to be measuring both hedonic and eudaimonic well-being constructs independent from other similar measures of well-being. For example, the WEMWBS and scales measuring hedonic wellbeing (e.g., PANAS-P, Watson et al., 1988; SWLS, Diener, et al., 1985) showed high correlations (supported by Tennant et al., 2007). Similarly, high correlations measuring eudaimonic well-being were found between WEMWBS and other measures of positive aspects of well-being (e.g., WHO-5, Bech, 2004; Psychological General Well-Being Index**,** PGWBI, Dupuy, 1984). The PGWBI measures psychological well-being (i.e., positive well-being, self-control).

In addition, the WEMWBS and negative affectivity constructs like the PANAS-N (The Positive and Negative Affect Schedule, Watson et al., 1988) and aspects of the PGWBI (Dupuy, 1984) (i.e., anxiety, depression) showed expected high negative correlations indicating good validity of the WEMWBS and that the WEMWBS is measuring separate constructs to other comparable measures. Similarly, Gremigni and Stewart-Brown (2011) reported a moderate negative correlation for the General Health Questionnaire (GHQ12;

[15] A bi-factorial model measures covariance between factors from an overarching factor, whilst a smaller subset of individual factors exists, accounting for variance (Rodriguez et al., 2015).

Rutter, 1997) (r = -.17) which measures an individual's perception of their current mental health. Finally, Stewart-Brown et al. (2009) investigated the criterion validity of the WEMWBS using Exploratory Factor Analysis. Findings showed a single underlying factor of well-being accounted for 54% of the variance, meaning that reliability of the tool was high. This included the GHQ12 (Rutter, 1997), where a statistically significant negative correlation (r = -.17) between the two measures was found findings which were similar to those found in the Tennant et al. (2007) study.

Construct validity

Construct validity is the extent to which a test measures what it proposes to measure (Kline, 1998). Tennant et al. (2007) conducted Confirmatory Factor Analysis for both samples (participant and population) on item responses to assess how well the structural equation models fit, in line with the expected one-factor construct of the WEMWBS. The analysis applied a goodness of fit[16] of indexes analysis (Goodness of Fit (GFI), Adjusted Goodness of Fit (AGFI), and Root Mean Square Error of Approximation (RMSEA)) for both samples (Tennant et al., 2007). The GFI and AGFI showed to be above required levels (GFI = 0.91 and AGFI = 0.87).The RMSEA fell below the upper limit, however it is suggested that this may have been related to the large sample size (Tennant et al., 2007). These findings demonstrate that both groups showed support for a one-factor scale format (questions loaded > 0.5 onto the single factor) (Tennant et al., 2007). The findings indicate that the WEMWBS has good construct validity (i.e., that it measures mental well-being). This an expected/favourable result. In addition, Tennant et al. (2007) set out to test the construct validity of the WEMWBS by looking for correlations between WEMWBS scores and factors

[16] Goodness of Fit refers to a statistical model that shows how well it fits a given set of observations. Measures of goodness of fit look at the difference between observed values and the expected values of a given model, which affects the reliability of a given measure (Kéry & Royle, 2016).

believed to be related to poor mental health (e.g., gender, socio-economic status); high scores would indicate that the measure had high validity and suggested that itis able to distinguish between constructs that contribute to good mental well-being and those that may lead to poor mental health well-being. As such, high scores on the WEMWBS should correlate with things that the literature has identified as contributing to good self-perceived levels of well-being. It was hypothesised that: men would obtain a higher score than women; there would be no relationship with age in relation to leaving education; there would be a positive relationship with higher socio-economic status. Research suggests that quality of life (in terms of status, stability, successfulness) is associated with higher mental well-being (Slade, 2010). As hypothesised, within the population sample (N = 2075), the WEMWBS score was significantly higher for men compared to women. Interestingly, higher levels of perceived wellbeing have been found for men compared to women (Tennant et al. (2007). Further, higher scores were shown for people aged 16–24 compared to people aged 55–74, suggesting that younger people have better wellbeing levels. Individuals who had completed their education aged 19 years or older had the highest mental well-being levels (Tennant et al., 2007).

Socio-economic status was also positively correlated with WEMWBS scores for both income levels and primary income provider social grade[17] (Tennant et al., 2007). In addition, higher scores were shown for people that owned and resided in their properties, and lower scores were obtained from widowed, divorced/separated, or unemployed individuals.

Stewart-Brown et al. (2009) investigated construct validity using a Scottish sample (N = 779). They reported that initially some items did not appear to be a good fit (i.e., “I’ve been interested in new things”). As a result, these items were removed, leaving a seven item scale

[17] A primary income provider refers to a member of a family who supports the family by being the main breadwinner (Bitton et al., 2019).

(SWEMWBS). The authors concluded that this shorter version was favourable to the full version of the WEMWBS in assessing mental well-being in populations. It was noted, however, that the shorter version loses aspects of hedonic well-being but does maintain a more complete picture of aspects of eudemonic wellbeing. Therefore, the SWEMWBS may represent a narrower understanding of mental well-being as a whole.

Content validity

Content validity refers to the degree to which a measure covers all areas of the construct in question. In reference to test development, this type of validity is concerned with whether the items in a measure are representative of the topic concerned (Kline, 2000). One way that content validity can be assessed is by determining the extent to which a measure is based on empirical evidence and expertise. In terms of the development of the WEMWBS, input was sought from a panel of experts (including professionals working in mental health settings) through a series of focus groups. The aim was to draw on their knowledge and experience in order to design a valid and robust measure of positively worded questions underpinned by a framework of mental well-being and psychological literature based on hedonic and eudaimonic factors (Tennant et al., 2007). However, it is important to consider that definitions of well-being may vary between cultures/language and are developed based on a certain framework/specific concept, as such they will be constrained to that specific definition used.

Further aspects of content validity include "the clarity of the instructions, linguistic aspects of the items (e.g., content, grammar), representativeness of the item pool, and the adequacy of the response format" (Koller et al., 2017, p. 1). Tennant et al. (2007) investigated the content validity of the WEMWBS by measuring the prevalence of complete answers and how widely distributed the responses to each item were. They investigated how relevant the

WEMWBS was for the target population by examining the trends of responses from both samples (student and Scottish population sample). Complete responses versus partial/omitted responses were compared using demographic information. Missing item responses were looked at to assess relevance and indicators of inappropriateness. However, the distribution of scores for both samples showed to be within a normal range, and the scale showed no ceiling effects. similar to other populations surveys, the WEMWBS was consistent with outcomes of other population surveys.

Furthermore, the frequency of the highest selected responses (e.g., 5 out 5) was analysed for both samples by looking at the distributions of responses from complete responders. For both samples, this showed few skewed distributions; further showing that all response options were used for all questions by at least one responder (Tennant et al., 2007). This illustrates that the target population understood the tool items suggesting that the construct of well-being, as defined by the tool, is being measured. A further consideration in respect of assessing content validity relates to the definition of mental well-being. As previously mentioned, there is disagreement as to whether purpose of life and spirituality should be included in a measure of well-being (Stewart-Brown et al., 2011). The authors of the WEMWBS chose not to include this construct as they did not feel there was adequate empirical evidence to support the inclusion of these items. A further rationale for not including these items was that concepts such as spirituality are posited to be culturally related; having more meaning and existence in some cultures, than in others (Stewart-Brown et al., 2011). For instance, a cross-cultural evaluation of the WEMWBS using focus groups found that Pakistani males and females in the UK reported a spiritual interpretation of mental well-being (Taggart et al., 2013) suggesting that the construct of mental well-being may not be generalisable for measurement across different cultural groups. Moreover, the study by Taggart et al. (2013) also explored minority ethnic participants' interpretations of the

questions in the WEMWBS. English speaking Chinese and Pakistani participants in the UK were sampled and focus groups were conducted. It was reported that Pakistani women struggled with the word *optimist*, as there is no equivalent translation in their native language. Younger Chinese and Pakistani males mistook *interest in others* as having sexual connotations. Furthermore,) Newbigging (2008, as cited in Taggart et al., 2013) reported that Chinese individuals understood the term *happiness* more than the concept of mental well-being, and Pakistani individuals considered mental well-being to be akin to peace of mind/contentment. Nonetheless, Taggart et al. (2013) concluded that all items were relatively straightforward to understand, and that items on the WEMWBS were transparent for minority populations. However, it is clear that further research needs to be conducted in this area to support claims that the measure is appropriate for minority populations as the subjective understanding of mental wellbeing may differ between cultural groups.

Face Validity

Face validity is concerned with how a test appears to the responder. Researchers purport that for face validity to be achieved, the purpose of a test should be clear to the responder (Kline, 2000). Face validity can have benefits, for example, participants may be more likely to complete a test if it is easy for them to see the purpose of the measure. However, where the purpose of the measure is explicit, there is an increased risk of false responses being provided by participants (Kline, 1998) due to response bias (i.e., a tendency for test takers to respond inaccurately). One such form of response bias is social desirability bias, where a respondent provides the answers that they believe are socially desirable/acceptable as opposed to an answer that depicts their genuine state of well-being. Tennant et al. (2007) assessed social desirability bias using the Balanced Inventory of Desirable Response (BIDR; Paulhus & Reid, 1991) with a sample of students in the UK. The

two sub-scales included were *impression management* and *self-deception*. Correlations with the sub-scales were lower than or similar to other measures of well-being (i.e., SWLS, Diener et al., 1985; WHO-5, Bech, 2004; PANAS-PA and PANAS-NA, Watson et al., 1988; and single-item GLS, Eurobarometer Life Satisfaction [http://ec.europa.eu/ public_opinion]).The authors concluded that the WEMWBS has good face validity and there did not appear to be concerns regarding response bias (Tennant et al., 2007). Both Stewart-Brown et al. (2009) and Gremigni and Stewart-Brown (2011) reported that social desirability bias for the original version of the WEMWBS was lower/equal to other comparable well-being scales.

Furthermore, in conducting the Italian validation of the WEMWBS, Gremigni and Stewart-Brown (2011) reported no concerns regarding social desirability. The research outlined here indicates that the WEMWBS has acceptable face validity within UK population samples of students, health professionals, and governance professionals. It is concluded that existing research has overall presented good face validity data for the original version of the tool.

Standardisation and Norms

Norms are usually the baseline or original sample from which a tool was developed, following which it is referred to as the comparison population or average against which subsequent sample scores can be compared (Kline, 2000). Norms are necessary in psychological research as they provide robust comparable groups of target populations. In short, "test norms enable determining the position of an individual test taker in the group." (Oosterhuis et al., 2015, p. 191). Norms allow expectations and generalisations to be made regarding a particular group of society but without a robust study using a sample that is representative of the population (e.g., including a range of genders, ages, and ethnicities) it is not possible to have accurate norms. When looking at norms, it is necessary to consider the size of the sample being used to calculate the norm, and how well it represents the target

population (Kline, 2000). Kline (2000) stipulates a sample size of more than 500 individuals as appropriate to minimise standard errors. The studies detailed in this review all exceeded sample sizes of more than 500 participants. However, it is of note that some of these studies used student samples (Clarke et al., 2011); as such it is argued that they may not be a representative sample of the general population (Bass et al., 2016; Gremigni & Stewart-Brown, 2011). For example, half of the sample used in Gremigni and Stewart-Brown's (2011) study were aged between 16 and 34 years, with their total sample failing to be equally representative of older age groups. However, Lloyd and Devine (2012) highlighted that their sample had a normal distribution in the general population, with low ceiling and floor effects. Lloyd and Devine (2012) further reported that Tennant et al.'s (2007) validation of the WEMWBS excluded two regions of the UK (Wales and Northern Ireland). To address this gap, Lloyd and Devine (2012) assessed the psychometric properties of the WEMWBS across a Northern Ireland sample and reported similar findings to Tennant et al. (2007). As the majority of studies focus on a student population, more studies need to be conducted to derive norms from a sample that is more representative of the general population.

Conclusions

This review has provided an overview of the development of the WEMWBS and discussed the measure in reference to its statistical reliability and validity.

The UK validation of the WEMWBS shows it to have some good psychometric characteristics (i.e., construct validity, content validity, and test-retest reliability). The research indicates that items from the WEMWBS map well onto the constructs of mental well-being that it seeks to measure (i.e., eudemonic and hedonic constructs; Gremigni & Stewart-Brown, 2011). The studies in this review have shown similar single factor (Ikink et al., 2012) or bi-factor (Gremigni & Stewart-Brown, 2011) structure (hedonic and eudaimonic

aspects) of the WEMWBS for a range of UK and other language adaptations. The converted Italian and Spanish versions of the WEMWBS showed good psychometric properties similar to the UK scale, demonstrating that adapted versions can be used with these specific populations (Lau et al., 2005). Although issues were highlighted for non-European cultural groups, where the definition of happiness is subjectively different, this could impact on the generalisability of the tool and its validity cross-culturally. Furthermore, in terms of positive findings, Gremigni and Stewart-Brown (2011) reported that the WEMWBS is a robust measure of mental well-being when applied at one week intervals with different groups. In addition, the review has shown the WEMWBS to have a robust evidence base, with good criterion validity. The WEMWBS was shown to positively correlate with similar measures assessing similar hedonic measures (i.e., The Positive and Negative Affect Schedule (PANAS-P), Watson et al., 1988) and negatively correlate with the eudemonic measures (i.e., The Positive and Negative Affect Schedule, (PANAS-N), Watson et al., 1988) (Gremigni & Stewart-Brown, 2011). However, it was of note that the WEBWBS lacked items that represented social well-being. Furthermore, lower correlations were found between scores on the WEBWBS and the Global Life Satisfaction (GLS; Diener et al., 1985), and Emotional Intelligence Scale (EIS; Malouff et al., 1998). These were used for comparative test purposes to ensure the WEBWBS was measuring a well-being construct. Low correlations indicate that variables have low relation to one another.

Some research suggests there to be few concerns regarding social desirability (Gremigni & Stewart-Brown, 2011; Stewart-Brown et al., 2009). Tennant et al. (2007) concluded that the WEMWBS has a sole focus on positive aspects of mental health; a good quick measure of mental well-being. No ceiling effects were found in the population sample. High correlations with other mental health and well-being scales were found.

Tennant et al. (2007) noted that testing for criterion validity (i.e., whether the outcome of the measure correlates with the outcomes of other measures of the same construct) of all of the comparable scales was not always possible. Therefore, it is recommended that further criterion validity tests are conducted. In addition, with reference to reliability, Taggart et al. (2013) reported high reliability of the measure with Chinese and Pakistani populations when completing the WEMWBS. The authors further reported good comprehension of the WEMWBS and that nearly all questions were clearly understood. However, the authors noted that cultural differences in the comprehension of mental well-being, both between and within communities, were present. Further validation studies are required to determine other components that may be relevant to the construct of well-being cross-culturally. Similar findings have been cited in studies looking at the validity of the WEMWBS in other countries (Castellví et al., 2013), again with population samples (Ikink et al., 2012; Maheswaran et al., 2012).

In terms of the internal consistency of the WEMWBS, it has been discussed in this critique that high internal consistency indicates different items on a scale all measure the same construct. However, very high levels of internal consistency could show some redundancy in items, termed *bloated specifics* (Oltmanns & Widiger, 2016). As noted above, findings regarding internal consistency are mixed and more research is required. A further point of consideration is the test-retest findings discussed in this review. Nearly all of the studies described a one-week time point for test re-test. It is therefore suggested that a longer time period between testing be implemented to test the stability of the WEMWBS over a longer period of time (Tennant et al., 2007).

In summary, despite the identified limitations, the WEMWBS is considered to be an appropriate tool for assessing mental well-being in UK samples (Tennant et al., 2007) as well

as Spanish, Italian, English speaking Chinese and Pakistani. The WEMWBS has been reported to be straightforward to complete and user friendly for test administrators, as demonstrated by its extensive use in clinical and research arenas. However, it is of note that there is a paucity of research validating the WEMWBS and a significant amount of this has been carried out by the authors who developed the measure meaning that there is a potential for researcher bias. It is recommended that the measure be further scrutinised and that the aforementioned limitations are borne in mind when interpreting an individual's score on the measure.

CHAPTER 5

Thesis Summary

Broadly speaking, this thesis aimed to add to the knowledge base around the use of Drama Based Therapies (DbTs) with forensic populations. As outlined in Chapter 1, the thesis aimed to: explore research on the impact of DbTs with forensic populations; gain insight into drama practitioner perspectives and experiences of a specific type of DbT; and critique a tool which has been used in the evaluation of a DbT intervention. These aims have been achieved through conducting three key pieces of work that are reported in chapters 2, 3, and 4 of this thesis. Chapter 2 comprised of a systematic review of the literature regarding the impact of DbT with forensic populations. Chapter 3 reported themes and sub-themes derived from interviews with current or former Geese Theatre Company actor practitioners, and Chapter 4 provides a critique of The Warwick-Edinburgh Mental Well-Being Scale (Tennant et al., 2006).

A summary of the main findings of each chapter will be provided below. In addition, due to the practice-based nature of the content of the thesis, implications for practice will be discussed. Lastly, some key strengths and limitations of the thesis will be provided, and recommendations made for future research directions.

Summary of findings

Chapter 2

The systematic literature review presented in Chapter 2 aimed to explore and synthesise research regarding the use of DbT with individuals in forensic settings. The review provided a short summary of types of DbT being used in forensic settings and provided a narrative synthesis of findings of qualitative and quantitative studies which explored the impact of a DbT used in a forensic setting.

Ten articles were included in the review. This comprised of four quantitative papers (Blacker, Watson, & Beech, 2008; Keulen-De Vos, et al., 2017; Koiv, & Kaudne, 2015; Reiss, et al., 1998), four qualitative papers (Colquhoun, Lord & Bacon, 2018; Stahler, 2007;

Gordon, Shenar, & Pendzik, 2018; Mundt, Marín, Gabrysch, Sepúlveda, Roumeau, & Heritage, 2019), and two mixed-method papers (Harkins et al., 2011; Stephenson & Watson, 2018). Half of studies included in the review investigated the impact of a DbT on levels of anger. Other issues explicitly addressed by DbTs were communication/relationship skills, problem solving, coping strategies, and confidence.

Despite methodological limitations of the articles included in the review, findings were promising regarding the positive impact that DbTs may have on issues such as anger/aggression, hope for the future, mental health/well-being, communication, gaining insight, and a sense of achievement.

The review also highlighted that a broad range of DbTs are being used in forensic settings, i.e., companies delivering DbT differed in their content an approach. However, there were common elements in approaches, such as the use of role-play and reflection. Furthermore, although the focus of the DbTs was not necessarily to modify a specific type of offending behaviour, issues such as communication, hope and mental well-being could be seen as being linked to an individual being less likely to re-offend (Prochaska & DiClemente, 1983; Ward et al., 2004).

It was noted that DbTs have an advantage over manualised Offender Behaviour Programmes in that they are less structured and are therefore able to be tailored to the needs of the individuals. Lastly, it was noted that research exploring the experiences and views of individuals who deliver DbTs would be a useful addition to the research base.

Chapter 3

The aim of chapter 3 was to explore the views and experiences of individuals who had delivered drama based herapy with forensic populations. Interviews were conducted with former and current Geese Theatre Company (GTC) Actor Practitioners (APs) who are/had delivered a programme to individuals in forensic settings. Previous research in this field has

shown encouraging findings regarding the effectiveness of DbTs with forensic populations in terms of helping to reduce maladaptive behaviours. The study therefore aimed to further these findings by gaining a wider understanding regarding, for example, what elements of the programs the APs felt were useful to forensic populations, strategies used to engage individuals, the value of the therapeutic relationship, the training they received, and challenges APs experienced in their role.

Thematic Analysis was used to analyse the transcripts and create themes and sub-themes from the data. The four key themes identified were: *The client journey*; *Actor Practitioner qualities and skills; Creativity vs manualised interventions; Challenges and experiences. The client journey* theme noted the importance of the process of behavioural change in those attending the DbT. APs referred to their experience of barriers to clients engaging with interventions (e.g., negative behaviours, and a lack of trust). These observations aligned with literature around the process of behavioural change (e.g., Prochaska & DiClemente, 1983) and the need to address barriers to change in order to increase engagement with a programme (Burrowes & Needs, 2009; Ward et al., 2004). In addition to the importance placed on building trust between themselves and prisoners, participants highlighted the importance of ensuring prisoners view the group as a safe place.

–As reported in the theme *Actor Practitioner (AP) qualities*, qualities which participants felt were necessary to possess in their role were, for example, acting skills (from past training and experience), an ability to empathise, and being able to build a rapport with individuals (e.g., by being authentic, finding common ground, and showing a genuine interest in them). These findings were consistent with literature that highlights the importance of building a good therapeutic relationship (Kozar & Day, 2012). Participants further noted that they valued the level of support and training they received throughout their time at GTC.

The third theme - *creativity vs manualised interventions* – summarised APs views regarding the flexibility and freedom the GTC afforded them in the delivery of programmes, i.e., not using a manual based intervention approach. APs further spoke about the use of masks in GTC's particular form of DbTs, and how unique this method is. Highlighting that they are a 'signature tool' with the GTC. APs expressed that through mask work people can develop and watch a character that they identify with which provides them with an objective view of themselves. Lastly, in theme four, APs discussed the challenges that the role involved, for example, the challenge of listening to difficult/distressing stories. However, they also noted that they found it to be a rewarding role. APs highlighted the importance of training and support within the team. In summary, the findings from the study supported that of previous literature regarding the positive impact of DbT and furthered such findings by providing a practitioner perspective.

Chapter 4

The critique presented in Chapter 4 aimed to explore the reliability and validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS; Stewart-Brown et al., 2007)

The findings showed that the WEMWBS has good psychometric characteristics (i.e., construct validity, content validity, and test-retest reliability), and – in line with previous research - the WEMWBS items map well onto the constructs of mental well-being (eudemonic and hedonic) (Gremigni & Stewart-Brown, 2011).

The studies in the review showed either single factor (Ikink et al., 2012) or bi-factor (Gremigni & Stewart-Brown, 2011) structure (hedonic and eudemonic aspects) of the WEMWBS for a range of UK and other language adaptations (i.e., Spanish, Italian, Chinese and Pakistan) (Castellví et al., 2013; Ikink et al., 2012; Maheswaran et al., 2012; Lau et al., 2005). However, discrepancies were found for non-European cultural groups, where the definition of happiness was found to be subjectively different. Taggart et al. (2013) noted that

cultural differences in the understanding of mental well-being with Chinese and Pakistani (English speaking) populations were present.

Gremigni and Stewart-Brown, (2011) reported that the WEMWBS is a robust measure of mental well-being when applied at one week intervals with different groups. Thus, showing good test retest reliability. However, it is of note that many of the studies in this review used a one-week time point for test re-test purposes. To ensure the stability of the WEMWBS over a longer period of time (Tennant et al., 2007) further testing was recommended.

Findings relating to the internal consistency of the WEMWBS were mixed. High levels of internal consistency were found (Tennant et al. (2007), which may suggest redundancy in items, a term called *bloated specifics* (Oltmanns & Widiger, 2016). The WEBWBS was further found to lack items representing social well-being (Tennant et al. (2007). Regarding social desirability, mixed findings were also noted (Gremigni & Stewart-Brown, 2011; Stewart-Brown et al., 2009), thus, warranting further research. Tennant et al. (2007) noted that testing for criterion validity of all of the comparable scales was not always possible. Therefore, further criterion validity tests are recommended.

In summary, the WEMWBS is considered to be a largely reliable and valid measure for assessing mental well-being in UK and other samples (Tennant et al., 2007). The review has shown the WEMWBS is straightforward to complete, and user friendly for test administrators. However, due to some shortcomings, an individual's results from the scale must be viewed with caution and it is suggested that additional measures of well-being are used.

Implications for practice

Each chapter of the thesis has made reference the potential practical implications of findings. It is noted that the findings may be of interest/use to practitioners working with forensic populations, in particular, those who use DbTs or techniques that would fall under

this banner. In addition, the findings may be of interest to academics when researching and teaching on the topic of 'what works' in the rehabilitation of offenders. More specifically, it is suggested that the following be considered by practitioners who deliver DbTs to forensic populations and other professionals/staff who work in a rehabilitative context with individuals who have committed a criminal offence:

- The importance of breaking down barriers to engagement (e.g., ice breaker exercises, group work, DbT techniques (e.g., mask work), creating a safe space)
- The value of a good Therapeutic relationship. Findings of Chapter 3 highlight certain qualities that can help build this relationship – e.g., empathy, being authentic, showing a genuine interest in each person, and finding common ground. In addition, the importance of spending time speaking with individuals 1:1 was highlighted.
- The importance of considering group dynamics within a treatment group; clients will benefit from getting to know and trust each other and build bonds.
- That interventions (both DbT and other rehabilitative interventions) need to consider the responsivity principle in their development and delivery. Interventions need to be responsive to the needs of individuals rather than adopting a 'one-size-fits-all' approach.
- There is a need for continued support and training for DbT practitioners. Practitioners may be exposed to distressing information. Reflective practice and supervision should be used effectively. Practitioners using DbT should possess acting skills to enhance the effectiveness of the intervention.
- Evaluations of DbT interventions should be conducted to further identify areas of good practice in DbT. Findings could be disseminated more widely in forensic services so that DbT techniques could be used within other interventions where appropriate.
- Scores on the WEMWBS should be viewed with caution and the measure should be used alongside other measures of mental well-being.

Recommendations for future research

Chapters 2, 3, and 4 of the thesis have included suggestions for future research. In sum, it is suggested that consideration is given to the range of measures which can be used to evaluate DbT interventions. For example, reconviction rates, behavioural reports and standardised self-report measures. Researchers should endeavour to use large sample sizes where possible to increase the generalisability of results. Longitudinal studies could be conducted to investigate the longer terms impacts of DbTs, i.e., whether the positive effects of taking part in a DbT are long lasting. Further research should be conducted to investigate the reliability and validity of measures used to evaluate DbTs, e.g., the WEMWBS.

Strengths and limitations

This thesis has contributed to the field of drama based therapy by synthesising relevant literature, exploring the views of Actor Practitioners, and critiquing a measure used in the evaluation of DbT. A strength of the literature review was that it included both qualitative and quantitative studies; thus, no studies were omitted based on research design. Although the DbTs differed in terms of approach and content, the review provided some detail regarding the DbT being evaluated in each paper. However, it is acknowledged that the papers did not contain highly detailed information about the particular intervention given that the focus was on the DbT evaluation, as such, it was not possible to provide the reader with detailed information and look at the effectiveness of specific techniques being used across all DbTs. In addition, due to the range of DbTs evaluated across all studies and small sample sizes, it is hard to generalise findings to other types of DbTs and other populations. A strength of the review was that there were clear inclusion/exclusion criteria which allowed for the identification of relevant papers. However, a limitation was that only studies using the English language were included which may have introduced bias.

The study presented in Chapter 3 was the first of its kind to explore a DbT from the perspective of individuals who deliver the GTC intervention. Valuable insights were provided regarding, for example, their views on how their specific type of DbT can impact on people, what methods they feel are effective, what qualities they feel are needed to facilitate the DbT, and what challenges they face. Unfortunately, due to the Covid-19 pandemic the sample included more former APs than current APs. Some participants hadn't worked for GTC for a number of months/years which may have led to inaccuracies in recall regarding the role. Further, when deriving themes from data the potential for researcher bias needs to be considered. Lastly, as the study was qualitative it is not possible to generalise results. For additional strengths and limitations please see discussion sections of Chapters 2 and 3.

Conclusions

This thesis has highlighted that a wide range of DbTs are being used with forensic populations. The key consistent finding across both the systematic literature review and the research project is that DbTs can have a positive impact on forensic populations. Although recidivism rates have not been used as an outcome measure, positive trends in terms of, for example, lower levels of aggression and improvements in mental well-being and communication skills, would indicate that DbTs could potentially decrease the likelihood of recidivism (Andrews & Bonta, 2010). However, further research would be needed to support this assumption. The thesis has also highlighted that there is flexibility within DbTs, i.e., the content and delivery can be tailored to meet the needs of the individual. It is hoped that the findings of the thesis will be of encouragement to those using DbT with forensic populations. It is also hoped that the findings will increase the profile of this form of intervention to encourage it being used more widely in forensic settings.

REFERENCES

Andrews, D. A., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Psychology Public Policy and Law*, *16* (1), 39–55. https://psycnet.apa.org/record/2010-01480-002.

Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct* (5th ed.). New Providence, New Jersey: LexisNexis Matthew Bender.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2011). The risk-need-responsivity (RNR) model: Does adding the good lives model contribute to effective crime prevention? Criminal Justice and Behavior, 38, 735–755. doi:10.1177/ 0093854811406356.

Antonowicz, D. H., & Ross, R. R. (1994). Essential components of successful rehabilitation programs for offenders. *International Journal of Offender Therapy and Comparative Criminology, 38* (2), 97–104. https://doi.org/10.1177/0306624X9403800202.

Armstrong, D., Gosling, A., & Weinman, J. (1997). The Place of Inter-Rater Reliability in Qualitative Research: An Empirical Study. Sociology-the Journal of The British Sociological Association, *Sociology.* 31. 597-606. https://journals.sagepub.com/doi/10.1177/0038038597031003015.

Baim, C., Brookes, S., & Mountford, A. (2002). *The Geese Theatre Handbook: Drama with Offenders and People at Risk.* Waterside Press.

Baim, C., Burmeister, J., & Maciel, M. (2007). *Psychodrama: Advances in theory and practice.* USA: Taylor and Frances.

Balakumar, P., Inamdar, M.N., & Jagadeesh, G. (2013). The Critical Steps For Successful Research: The Research Proposal And Scientific Writing: (A Report On The Pre-Conference Workshop Held In Conjunction With The 64th Annual Conference Of The Indian Pharmaceutical Congress-2012). *The Journal of Pharmacology & Pharmacotherapeutics,* 4 (2), 130–138. https://pubmed.ncbi.nlm.nih.gov/23761709/.

Baldwin, S.A., Wampold, B.E., & Imel, Z.E. (2007). Untangling the alliance-outcome correlation: exploring the relative importance of therapist and patient variability in the alliance. *Journal of Consulting and Clinical Psychology, 75*, 842-852. https://psycnet.apa.org/record/2007-19013-002.

Bass, M., Dawkin, M., Muncer, S., Vigurs, S., & Bostock, J. (2016). Validation of Warwick-Edinburgh Mental Well-being Scale (WEMWBS) in a population of people using Secondary Care Mental Health Services. *Journal of Mental Health, 25* (4), 323-329. https://www.tandfonline.com/doi/full/10.3109/09638237.2015.1124401.

Batcup, D. C. (2013). A discussion of the dance movement psychotherapy literature relative to prisons and medium secure units. *Body, Movement and Dance in Psychotherapy*, 8 (1), 5–16. https://www.tandfonline.com/doi/abs/10.1080/17432979.2012.693895.

Bech, P. (2004). WHO Wellbeing Index: Measuring the dimensions of psychological general well-being by the WHO-5. *QoL Newsletter*, *32*, 15-16. https://eprovide.mapi-trust.org/instruments/who-five-well-being-index.

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The Hopelessness Scale. *Journal of Consulting and Clinical Psychology, 42* (6), 861–865. https://doi.org/10.1037/h0037562.

Belur, J., Tompson, L., Thornton, A., & Simon, M. (2021). Interrater Reliability in Systematic Review Methodology: Exploring Variation in Coder Decision-Making. *Sociological Methods & Research*, *50*(2), 837–865. https://doi.org/10.1177/0049124118799372.

Bergman, J. (2000). Creating new cultures: Using drama therapy to build therapeutic communities in prisons. In P. Lewis, & D. R. Johnson (Eds.), *Current approaches in drama therapy.* Springfield, Illinois, US: Charles C Thomas Publisher Ltd.

Bergman, J. (2017). The roar of the 1960s and why we chose to work inside the system. *Drama Therapy Review Hollins University and Geese Theatre Company*, 3 (2). 10.1386/dtr.3.2.321_7.

Bitton, A., Fifield, J., Ratcliffe, H., Karlage, A., Wang, H., Veillard, J. H., Schwarz, D., & Hirschhorn, L. R. (2019). Primary healthcare system performance in low-income and middle-income countries: a scoping review of the evidence from 2010 to 2017. *BMJ global health*, *4* (Suppl 8), e001551. https://doi.org/10.1136/bmjgh-2019-001551.

Blacker, J., Watson, A., & Beech, A. R. (2008). A Combined Drama based And CBT Approach To Working With Self-Reported Anger And Aggression. *Criminal Behaviour And Mental Health, 18*, 129-137. https://pubmed.ncbi.nlm.nih.gov/18383198/.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical methods in medical research*, *8*(2), 135-160. https://pubmed.ncbi.nlm.nih.gov/10501650/.

Bourgon, G., & Guiterrez, L. (2013). *The Importance of Building Good Relationships in Community Corrections: Evidence, Theory and Practice of the Therapeutic Alliance.*

In: Ugwudike P., Raynor P. (eds) What Works in Offender Compliance. Palgrave Macmillan, London. https://doi.org/10.1057/9781137019523_15

Bragg-Levanas, D. (2014). The Institutionalized Body: An Investigation Into The Use Of Drama Therapy On A Forensic Psychiatric Unit. Submitted In Partial Fulfillment Of The Requirements For The Degree Of Master Of Arts In Drama Therapy At The Steinhardt School Of Culture, Education, And Human Development. New York University. https://www.academia.edu/20859779/The_institutionalized_body_An_investigation_into_the_use_of_drama_therapy_on_a_forensic_psychiatric_unit.

Braun, V. & Clarke, V. (2006). Using Thematic Analysis In Psychology. *Qualitative Research In Psychology,* 3 (2), 77-101. https://www.researchgate.net/publication/235356393_Using_thematic_analysis_in_psychology.

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health, 11*. 1-9. 10.1080/2159676X.2019.1628806.

Braunholtz, S., Davidson, S., Myant, K., & O'Connor, R. (2006). *Well? What do you think? The third national Scottish survey of public attitudes to mental health, mental wellbeing and mental health problems.* Scottish Executive: Edinburgh. https://hqlo.biomedcentral.com/articles/10.1186/1477-7525-5-63.

Brooks, R., Robin, R., & de Charro, F. (2003). EuroQol Health Status Visual Analogue Scale. (Eds.): *The measurement and valuation of health status using EQ5-D: a European perspective.* Kluwer Academic Publishers.

Burrowes, N., & Needs, A. (2009). Time to contemplate change? A new framework for assessing readiness to change with offenders. *Aggression and Violent Behaviour,* 14, 39– 49. https://www.researchgate.net/publication/223209655_Time_to_contemplate_change_A_framework_for_assessing_readiness_to_change_with_offenders.

Carruthers, Cynthia & Hood, Colleen. (2004). The power of the positive: Leisure and wellbeing. Therapeutic recreation journal. 38. 225-245.

Cassidy, S. (2013). An Exploration Of The Processes That Underlie Change In Dramatherapy: A Grounded Theory Analysis. D Clin Psy Thesis, University Of Glasgow. Http://Theses.Gla.Ac.Uk/4611.

Cassidy, S., Gumley, A. & Turnbull, S. (2017). Safety, Play, Enablement, And Active Involvement: Themes From A Grounded Theory Study Of Practitioner And Client Experiences Of Change Processes In Dramatherapy. *The Arts In Psychotherapy*, 55, 174–185. https://www.researchgate.net/publication/264393411_Exploring_core_processes_facilitating_therapeutic_change_in_Dramatherapy_A_grounded_theory_analysis_of_published_case_studies.

Casson, J. (2004). *Drama, Psychotherapy and Psychosis: Dramatherapy and Psychodrama with People Who Hear Voices.* Brunner-Routledge.

Castellví, P., Forero, C. G., Codony, M., Vilagut, G., Brugulat, P., Medina, A., Gabilondo, A., Mompart, A., Colom, J., Tresserras, R., Montserrat F., Stewart-Brown, S. & Alonso, J. (2013). The Spanish Version Of The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) Is Valid For Use In The General Population. *Quality of Life Research,* 1-12. https://www.researchgate.net/publication/256449294_The_Spanish_Version_of_the_

Warwick-Edinburgh_Mental_Well-Being_Scale_WEMWBS_is_Valid_for_Use_in_the_General_Population.

Cattell, R. B. (1973). *Personality and mood by questionnaire*. London, UK: Jossey-Bass.

Chandler, M. (1973). Egocentrism and antisocial behaviour: The assessment and training of social perspective skills. *Developmental Psychology, 44*, 326-333. https://www.semanticscholar.org/paper/Egocentrism-and-antisocial-behavior%3A-The-assessment-Chandler/35c217857843f2832d8854fa534463848b8c9942.

Chesner, A. (1995). *Dramatherapy for people with learning disabilities: a world of difference.* London, UK: Jessica Kingsley.

Clarke, A., Putz, R., Friede, T., Ashdown, J., Adi, Y., Martin, S., Flynn, P., Blake, A., Stewart-Brown, S., & Platt, S. (2011). Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): Validated For Teenage School Participants In England And Scotland. A Mixed Methods Assessment. *BMC Public Health*, 11, 487. https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-11-487.

Clarke, V., & Braun, V. (2013). *Successful Qualitative Research: A Practical Guide for Beginners.*

Cognition & emotion, *21*(6), 1212–1237.

Colquhoun, B., Lord, A., & Bacon, A.M. (2018) A Qualitative Evaluation of Recovery Processes Experienced by Mentally Disordered Offenders Following a Group Treatment Program, *Journal of Forensic Psychology Research and Practice,* 18 (5), 352-373, 10.1080/24732850.2018.1510280. https://www.tandfonline.com/doi/abs/10.1080/24732850.2018.1510280.

Cooke, A., Smith, D., & Booth, A. (2012). Beyond PICO: The SPIDER Tool For Qualitative Evidence Synthesis. *Qualitative Health Research.* 22. 1435-43. https://www.researchgate.net/publication/230565751_Beyond_PICO_the_SPIDER_tool_for_qualitative_evidence_synthesis.

Critical Appraisal Skills Programme. (2017). CASP Qualitative Research Checklist. Retrieved From Http://Docs.Wixstatic.Com/Ugd/Dded87_25658615020e427da194a325e7773d42.Pdf

Cuijpers, P., Sijbrandij, M., Koole, S.L., Andersson, G., Beekman, A.T., & Reynolds, C.F. (2014). Adding Psychotherapy To Antidepressant Medication In Depression And Anxiety Disorders: A Meta-Analysis. *World Psychiatry,* 13 (1), 56–67. https://pubmed.ncbi.nlm.nih.gov/24497254/.

Cullen, F., & Gendreau, P. (2000). Assessing Correctional Rehabilitation: Policy, Practice, and Prospects. *Criminal Justice. 3.* https://www.researchgate.net/publication/241385469_Assessing_Correctional_Rehabilitation_Policy_Practice_and_Prospects.

Cullen, F., & Gendreau, P. (2001). From Nothing Works to What Works: Changing Professional Ideology in the 21st Century. *Prison Journal. 81.* 313-338. https://www.researchgate.net/publication/249707407_From_Nothing_Works_to_What_Works_Changing_Professional_Ideology_in_the_21st_Century.

Davey, L., Day, A., & Balfour, M. (2014). *Performing Desistance: How Might Theories of Desistance From Crime Help Us Understand the Possibilities of Prison Theatre.* https://doi.org/10.1177/0306624X14529728.

Day, A., Bryan, J., Davey, L., & Casey, S. (2006). The process of change in offender rehabilitation programmes. *Psychology, Crime and Law,* 12, 473-487. https://psycnet.apa.org/record/2006-20005-003.

Day, A., Howells, K., Casey, S., Ward, T., Chambers, J., & Birgden, A. (2009). Assessing treatment readiness in violent offenders. *Journal of Interpersonal Violence,* 24, 18-35. https://psycnet.apa.org/record/2013-22525-002.

Dehnavi, S., Bajelan, M., Pardeh, S.J., Khodaviren, H., & Dehnavi, Z. (2016). The Effectiveness Of Psychodrama In Improving Quality Of Life Among Opiate-Dependent Male Patients. *International Journal Of Medical Research & Health Sciences,* 5 (5), 243–7. https://sciexplore.ir/Documents/Details/192-233-588-244.

Diener, E., Emmons, R.A., Larsen, R.J., & Griffin, S. (1985). The Satisfaction with Life Scale (SWLS). *Journal of Personality Assessment*, *49* (1), 71-75. https://www.tandfonline.com/doi/abs/10.1207/s15327752jpa4901_13.

Dixon-Woods, M., Sutton, A. J., Shaw, R., Miller, T., Smith, J., Young, B., Bonas, S., Booth, A., & Jones, D. (2007). Appraising Qualitative Research For Inclusion In Systematic Reviews: A Quantitative And Qualitative Comparison Of Three Methods. *Journal Of Health Services Research & Policy,* 12 (1), 42-48. https://pubmed.ncbi.nlm.nih.gov/17244397/.

Downes, M. J., Brennan, M. L., Williams, H. C., & Dean, R. S. (2016). Development Of A Critical Appraisal Tool To Assess The Quality Of Cross-Sectional Studies (AXIS). *The British Medical Journal, Open,* 6 (12), 1-7. https://bmjopen.bmj.com/content/6/12/e011458.

Dupuy, H.J. (1984). The Psychological General Well-Being (PGWB) Index. In *Assessment of quality of life in clinical trials of cardiovascular therapies*. Edited by: Wenger N. New York: Le Jacq:*170–183.* https://eprovide.mapi-trust.org/instruments/psychological-general-well-being-index.

Farrant, B. (2012). *A Counselling psychology perspective on the experiences of therapists working with clients convicted of crime* [Unpublished doctoral thesis]. City University of London.

Fazel, S., & Seewald, K. (2012). Severe mental illness in 33 588 prisoners worldwide: Systematic review and meta-regression analysis. *British Journal of Psychiatry, 200* (5), 364–373. https://doi.org/10.1192/bjp.bp.111.096370.

Feniger -Schaal, R., & Orkibi, H. (2020). Integrative systematic review of drama therapy intervention research. *Psychology of Aesthetics, Creativity, and the Arts, 14*(1), 68–80. https://doi.org/10.1037/aca0000257.

Feniger-Schaal, R. (2016). A Dramatherapy Case Study With A Young Man Who Has Dual Diagnosis Of Intellectual Disability And Mental Health Problems. *The Arts In Psychotherapy*, 50, 40–45. https://psycnet.apa.org/record/2016-45484-007.

Fleeman, N., & Dundar, Y. (2014). Data Extraction: Where Do I Begin? In A. Boland, M. G. Cherry, & R. Dickson (Eds.), *Doing A Systematic Review: A Student's Guide* (Pp. 85-98). London, UK: Sage.

Foa, E. B., Keane, T. M., Friedman, M. J., & Cohen, J. A. (Eds.). (2009). *Effective treatments for PTSD: Practice guidelines from the International Society for Traumatic Stress Studies* (2nd ed.). The Guilford Press.

Friedrich, M., & Leiper, R. (2006). Countertransference Reactions in Therapeutic Work with Incestuous Sexual Abusers. *Journal Of Child Sexual Abuse*, *15.* 51 - 68. 10.1300/J070v15n01_03.

Frost, A., & Connolly, M. (2004). Reflexivity, Reflection, and the Change Process in Offender Work. *Sexual Abuse,* 16 (4), 365–380. https://doi.org/10.1177/107906320401600408.

Garcia, A., & Buchanan, D. R. (2009). Psychodrama. In D. R. Johnson & R. Emunah (Eds.), *Current approaches in drama therapy* (2nd ed., pp. 393–423). Springfield, IL: Charles C. Thomas.

Gersie, A. (1996). *Dramatic Approaches to Brief Therapy.* Jessica Kingsley Publishers.

Glaser, B., & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research.* Mill Valley, CA: Sociology Press.

Goldberg, D.P., & Williams, P. (1988). *A user's guide to the General Health Questionnaire.* Windsor, UK: NFER-Nelson.

Goodman R. (1997). The Strengths and Difficulties Questionnaire: a research note. Journal of child psychology and psychiatry, and allied disciplines, 38 (5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Gordon, J. Shenar, Y., & Pendzik, S. (2018). Clown Therapy: A Drama Therapy Approach To Addiction And Beyond. The Arts In Psychotherapy, 57, 88–94. https://www.researchgate.net/publication/321735013_Clown_therapy_A_drama_therapy_approach_to_addiction_and_beyond.

Gregoire, G., Derderian, F., & Le Lorier, J. (1995). Selecting the language of the publications included in a meta-analysis: Is there a tower of babel bias? *Journal of Clinical Epidemiology,* 48 (1), 159-163.

Gremigni, P., & Stewart-Brown, S. (2011). Measuring mental well-being: Italian validation of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS). *Giornale Italiano Di Psicologia, 38 (*2), 485-508. http://wrap.warwick.ac.uk/44112/.

Hackshaw, A. (2008). Small Studies: Strengths and Limitations. *European Respiratory Journal,* 32 (5). 1141-1143. https://erj.ersjournals.com/content/erj/32/5/1141.full.pdf.

Harkins, L., Pritchard, C., Haskayne, D., Watson, A., & Beech, A.R. (2011). Evaluation Of Geese Theatre's Re-Connect Program: Addressing Resettlement Issues In Prison. *International Journal Of Offender Therapy And Comparative Criminology, 55* (4), 546–66. https://journals.sagepub.com/doi/abs/10.1177/0306624X10370452. https://doi.org/10.1080/02699930701438020.

Hu, Y., Stewart-Brown, S., Twigg, L., & Weich, S. (2007). Can the 12 item General Health Questionnaire be used to measure positive mental health? *Psychological Medicine,* 37(7):1005-13. https://pubmed.ncbi.nlm.nih.gov/17274855/.

Hughes, J. (2005). *Doing the Arts Justice: A review of research literature, practice and theory.* London. https://www.researchgate.net/publication/237311511_Doing_the_Arts_Justice_A_Review_of_Research_Literature_Practice_and_Theory.

Ikink, J. G., Lamers, S. M., & Bolier, J. M. (2012). De Warwick-Edinburgh Mental Well-Being Scale (WEMWBS). *Als Meetinstrument Voor Mentaal Welbevinden In Nederland.* https://www.semanticscholar.org/paper/De-Warwick-Edinburgh-Mental-Well-being-Scale-als-in-Ikink/5bf5323eb837e805f58d027ae92d1d8d35c8b8a5.

Jennings, S., McGinley J., & Orr, M. (1997). Masking and unmasking: dramatherapy with offender patients. In Jennings S (ed.), *Dramatherapy: Theory and Practice 3.* London: Routledge.

Jerusalem, M., & Schwarzer, R. (1992). Self-efficacy as a resource factor in stress appraisal processes. In R. Schwarzer (Ed.), *Self-efficacy: Thought control of action* (p. 195–213). Hemisphere Publishing Corp.

Jones, P. (2007). *Drama as therapy: theory, practice, and research* (2nd ed.). London; New York: Routledge.

Joseph, S., Linley, P. A., Harwood, J., Lewis, C. A., & McCollam, P. (2004). Rapid assessment of well-being: The Short Depression-Happiness Scale (SDHS). *Psychology and psychotherapy,* 77 (Pt 4), 463–478. https://doi.org/10.1348/1476083042555406.

Joseph, S., Linley, P.A., Harwood, J., Lewis, C.A., & McCollam, P. (2004). Short Depression Happiness Scale. Rapid assessment of well-being: the Short Depression-Happiness Scale (SDHS). *Psychology and Psychotherapy: Theory, Research and Practice, 77,* 463 - 478. https://arizona.pure.elsevier.com/en/publications/rapid-assessment-of-well-being-the-short-depression-happiness-sca.

Journal of the British association for drama therapists, (1992). Dramatherapy, 15:1, 3, DOI: 10.1080/02630672.1992.9689830.

Kammann, R., & Flett, R. (1983). A Scale To Measure Current Level Of General Happiness. *Australia Psycholo*gy, *35*, 259–265. https://doi.org/10.1080/00049538308255070.

Kedem-Tahar, E., & Kellermann, N. (1996). Psychodrama and Drama Therapy: A Comparison. *Arts in Psychotherapy, 23,* 27-36. 10.1016/0197-4556(95)00059-3. https://www.semanticscholar.org/paper/Psychodrama-and-drama-therapy%3A-A-comparison-Kedem-Tahar-Felix-Kellermann/8164a77469540d395b9f28a7d816928fb8b32367.

Keisari, S., Yaniv, D., Palgi, Y., & Gesser-Edelsburg, A. (2018). Conducting Playback Theatre with Older Adults – A Therapist's Perspective. *The Arts in Psychotherapy, 60.* 10.1016/j.aip.2018.07.002.

Kéry, M. & Royle, J. A. (2016). Applied Hierarchical Modeling in Ecology.

Keulen-De Vos, M., Van Den Broek, E.P.A., Bernstein, D.P., Vallentin, R., & Arntz, A. (2017). Evoking Emotional States In Personality Disordered Offenders: An

Experimental Pilot Study Of Experiential Techniques. *The Arts In Psychotherapy,* 53, 80–88. https://psycnet.apa.org/record/2017-13328-011.

Kline, P. (1998). *The New Psychometrics*. London, UK: Routledge.

Kline, P. (2000). *Handbook Of Psychological Testing* (2nd Ed.). London, UK: Routledge.

Koch, S. C., Kunz, T., Lykou, S., & Cruz, R. F. (2014). Effects of dance movement therapy and dance on health-related psychological outcomes: A meta-analysis. *The Arts in Psychotherapy,* 41(1), 46–64. 10.1016/j.aip.2013.10.004.

Koiv, K. & Kaudne, L. (2015). Impact Of Integrated Arts Therapy: An Intervention Program For Young Female Offenders In Correctional Institution. *Psychology,* 6, 1-9. https://psycnet.apa.org/record/2015-06297-011.

Koller, I., Levenson, M. R., & Glück, J. (2017). What do you think you are measuring? A mixed-methods procedure for assessing the content validity of test items and theory-based scaling. *Frontiers in Psychology, 8,* 126. https://doi.org/10.3389/fpsyg.2017.00126.

Kozar, C., & Day, A. (2012). The therapeutic alliance in offending behaviour programs: A necessary and sufficient condition for change? *Aggression and Violent Behaviour, 17*. 482 - 87. 10.1016/j.avb.2012.07.004.

Landy R. (1994). *Drama therapy: Concepts and practices*. (2nd ed.). Springfield, Ill. : C.C. Thomas.

Lau, A. L. D., Cummins, R. A., & McPherson, W. (2005). An Investigation into the Cross-Cultural Equivalence of the Personal Well-being Index. *Social Indicators Research*, *72*, 403 - 432. https://link.springer.com/article/10.1007/s11205-004-0561-z.

Lavrakas, P.J. (2008) Encyclopedia of survey research methods. Sage Publications, Inc., Thousand Oaks. https://dx.doi.org/10.4135/9781412963947.

Leff, J., & Warner, R. (2006). *Social inclusion of people with mental illness*. New York: Cambridge University Press. doi:10.1017/CBO9780511543937.

Leite, A., Medeiros, A.G., Rolim, C., Pinheiro, K., Beilfuss, M., Leão, M., Castro, T., & Júnior, J.A. (2019). Hope Theory and Its Relation to Depression: A Systematic Review. *Annals of Psychiatry* and *Clinical Neuroscience*, 2 (2):1014. http://www.remedypublications.com/open-access/hope-theory-and-its-relation-to-depression-a-systematic-review-168.pdf.

Lessner Listiakova, I. (2015). *Tertiary education in expressive therapies in the European context. In book: International Context of Tertiary Education in Expressive Therapies,* (pp.12-35). Publisher: Univerzita Palackého v Olomouci.

Lessner Listiakova, I., & Valenta, M. (2015). Therapist self-evaluation instrument in dramatherapeutic process: a pilot study, (pp. 88-94). Conference: Proceedings of INTCESS15 - 2nd International Conference on Education and Social Sciences. https://www.researchgate.net/publication/280494940_Therapist_self-evaluation_instrument_in_dramatherapeutic_process_a_pilot_study.

Lessner Listiakova, I., & Valenta, M. & Müller, O., & Svoboda, P. (2014). Evaluation of dramatherapy process with clients of psychiatric hospital. International multidisciplinary scientific conference. *Social sciences & arts SGEM2014*, 167-174. 10.5593/sgemsocial2014/B11/S1.023. https://pure.northampton.ac.uk/en/publications/evaluation-of-dramatherapy-process-with-clients-of-psychiatric-ho.

Lim, C.R., Harris, K., Dawson, J., Beard, D.J., Fitzpatrick, R., & Price, A.J. (2015). Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. *BMJ Open.* 27; 5 (7) :e007765. 10.1136/bmjopen-2015-007765. PMID: 26216152; PMCID: PMC4521553.

Lloyd, K., & Devine, P. (2012). Psychometric Properties Of The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) In Northern Ireland. *Journal Of Mental Health,* 21, 257 – 63. https://pubmed.ncbi.nlm.nih.gov/22574955/.

Lofland, J., Snow, D., Anderson, L., & Lofland, L. (2006). *Analyzing social settings. A guide to qualitative observation and analysis.* Belmont, CA: Wadsworth/Thomson Learning.

Lopez, M. A., Gabilondo, A., Codony, M., Garcia-Forero, C., Vilagut, G., Castellvi, P., Ferrer, M., & Alonso, J. (2013). Adaptation into Spanish of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) and Preliminary Validation in a Student Sample, *Quality of Life Research, 22* (5),1099–104. https://link.springer.com/article/10.1007%2Fs11136-012-0238-z.

Lord, A., & Willmot, P. (2004). The process of overcoming denial in sexual offenders. *Journal of Sexual Aggression. 10*. 51-61. 10.1080/13552600410001670937.

Lowenkamp, C., Holsinger, A., Robinson, C.R., & Cullen, F.T. (2012). When a person isn't a data point: Making evidence-based practice work. *Federal Probation, 76.* 11-21. https://www.researchgate.net/publication/286896951.

Maas, J. (1966). The use of actional procedures in group psychotherapy with sociopathic women. *International Journal of Group Psychotherapy*, 16, 190-197. https://psycnet.apa.org/record/1966-08975-001.

Maheswaran, H., Weich, S., Powell, J., & Stewart-Brown, S. (2012). Evaluating The Responsiveness Of The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): Group And Individual Level Analysis. *Health Quality Life Outcomes, 10*, 156–64. https://hqlo.biomedcentral.com/articles/10.1186/1477-7525-10-156.

Matarasso, F. (1997). *Use or ornament? The social impact of participation in the arts*. Stroud, UK: Comedia. https://www.artshealthresources.org.uk/docs/use-or-ornament-the-social-impact-of-participation-in-the-arts/.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & emotion*, *23*(2), 209–237. https://doi.org/10.1080/02699930802204677.

Meekums, B., & Daniel, J. (2011). Arts with offenders: A literature synthesis. *The Arts in Psychotherapy, 38* (4), 229–238. https://research.edgehill.ac.uk/en/publications/arts-with-offenders-a-literature-synthesis.

Melnick, M. (1984). Skills through drama: The use of professional theatre techniques in the treatment and education of prison and ex-offender populations. *Journal of Group Psycho-therapy and Psychodrama, 37*, 104-116.

Mundt, A.P., Marín, P., Gabrysch, C., Sepúlveda, C., Roumeau, J., & Heritage, P. (2019). Initiating Change Of People With Criminal Justice Involvement Through Participation In A Drama Project: An Exploratory Study. *Psychiatry,* 10, 716. https://www.frontiersin.org/articles/10.3389/fpsyt.2019.00716/full.

NHS Health Scotland: Health Education Population Survey: Update from 2006 survey. *Glasgow: NHS Health Scotland (in preparation) [database on the Internet]* 2007 [http://www.data- archive.ac.uk]. Colchester: UK Data Archive. https://hqlo.biomedcentral.com/articles/10.1186/1477-7525-5-63.

North American Drama Therapy Association. (2014). *What is drama therapy?* http://www.nadta.org/what-is-drama-therapy.html [2014-10-10].

Novick, G. (2008). Is there a bias against telephone interviews in qualitative research? *Research Nursing Health, 31(4)*, 391-398. 10.1002/nur.20259

O'Connor, P., & O'Connor, B. (2009). Editorial: Research in Drama Education. *The Journal of Applied Theatre and Performance*, *14 (4)*, 471-477. https://doi.org/10.1080/13569780903285966.

Oosterhuis, H. E. M., van der Ark, L. A., & Sijtsma, K. (2016). Sample Size Requirements for Traditional and Regression-Based Norms. *Assessment*, *23* (2), 191–202. https://doi.org/10.1177/1073191115580638.

Papagiannaki, A., & Shinebourne, P. (2016). The contribution of creative art therapies to promoting mental health: An IPA study of therapists' understandings of working with self-stigmatisation. *The Arts in Psychotherapy, 50.* 10.1016/j.aip.2016.06.007.

Paulhus, D.L., & Reid, D.B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology, 60*, 307-317. https://psycnet.apa.org/record/1991-18338-001.

Pendzik, S. (2006). On dramatic reality and its therapeutic function in drama therapy. *The Arts in Psychotherapy, 33* (4), 271-280. https://psycnet.apa.org/record/2006-11253-001.

Phillips, L. (1997). *In the public interest: Making art that makes a difference in the United States*. Stroud, UK: Comedia.

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC Methods Programme. Http://Citeseerx.Ist.Psu.Edu/Viewdoc/Download?Doi=10.1.1.178.3100&Rep=Rep1 & Type=Pdf.

Pope, C., Mays, N., & Popay, J. (2008). Synthesizing Qualitative And Quantitative Health Evidence: A Guide To Methods. *Sociology for Health and Illness,* 30, (2), 330-331. https://onlinelibrary.wiley.com/doi/full/10.1111/j.1467-9566.2007.1077_5.x.

Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology, 51* (3), 390–395. https://doi.org/10.1037/0022-006X.51.3.390.

Ramsden, E., & Guarnieri, M. (2010) *Dramatherapy and victim empathy: a workshop approach in a forensic setting*. In, Jones, P., Drama as Therapy volume 2. UK, Routledge.

Rasch, G. (1977). On Specific Objectivity: An attempt at formalizing the request for generality and validity of scientific statements. *The Danish Yearbook of Philosophy*, 14, 58-93.

Ratner, B. (2009). The correlation coefficient: Its values range between +1/−1, or do they?. *Journal of Targeting Measurement* and *Analysis for Marketing. 17*, 139–142. https://doi.org/10.1057/jt.2009.5.

Reason, M. (2019). A prison audience: women prisoners, Shakespeare and spectatorship, *Cultural Trends, 28* (2-3), 86-102. 10.1080/09548963.2019.1617929.

Reiss, D., Quayle, M., Brett, T., & Meux, C. (1998). Dramatherapy for mentally disordered offenders: Changes in levels of anger. *Criminal Behaviour and Mental Health,* 8, 139-153. https://onlinelibrary.wiley.com/doi/abs/10.1002/cbm.232.

Revicki, D. (2014). Internal Consistency Reliability. *Encyclopedia of Quality of Life and Well-Being Research* (A. C. Michalos, Ed.). Springer. https://doi.org/10.1007/978-94-007-0753-5_1494.

Robson, C. (2002). *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*, (2nd ed.). Oxford: Blackwell Publishers Ltd.

Rodriguez, A., Reise, S, & Haviland, M. (2015). Applying Bifactor Statistical Indices in the Evaluation of Psychological Measures. *Journal of personality assessment.* 98. 1-15. 10.1080/00223891.2015.1089249.

Ross, A. O., Lacey, H. M., & Parton, D. A. (1965). The Development of a Behavior Checklist for Boys. *Child Development,* 36, 1013-1027. http://dx.doi.org/10.2307/1126940.

Ross, R. R., & Hilborn, J. (2008). *Rehabilitating rehabilitation: Neurocriminology for treatment of antisocial behavior.* Ottawa: Cognitive Centre of Canada. https://www.researchgate.net/publication/281625853_Rehabilitating_Rehabilitation_Neurocriminology_for_Treatment_Prevention_of_antisocial_Behavior.

Ross, R. R., Fabiano, E. A., & Ross, R. D. (1986). *Reasoning and Rehabilitation: A handbook for teaching cognitive skills.* Ottawa, Canada: Center for Cognitive Development. https://www.researchgate.net/publication/281625853_Rehabilitating_Rehabilitation_Neurocriminology_for_Treatment_Prevention_of_antisocial_Behavior.

Ruddy, R., & Dent-Brown, K. (2007). Drama Therapy For Schizophrenia Or Schizophrenia-Like Illnesses. *Cochrane Database Of Systematic Reviews,* 1, CD005378. 10.1002/14651858.CD005378.Pub2. https://pubmed.ncbi.nlm.nih.gov/17253555/.

Rutter, C. (1997). GHQ-12 General Health Questionnaire.: The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological*

Medicine, 27, 191-197. https://hqlo.biomedcentral.com/articles/10.1186/1477-7525-1-66.

Ryff, C. D., & Keyes, C. L. M. (1995). SPWB Scale of Psychological Well-Being: The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, *69*, 719-727. https://pubmed.ncbi.nlm.nih.gov/7473027/.

Salkind, N. J. (2010). *Encyclopedia of research design* (Vols. 10). Thousand Oaks, CA: SAGE Publications, Inc. 10.4135/9781412961288.

Sapouna, M., Bisset, C., Conlong, A., & Matthews, B. (2015). *What Works to Reduce Reoffending: A Summary of the Evidence. Edinburgh: Scottish Government.*

Sarre, R. (2001). Beyond 'What Works?' A 25-year Jubilee Retrospective of Robert Martinsons Famous Article. *Australian and New Zealand Journal of Criminology. 34,* 38-46. https://journals.sagepub.com/doi/abs/10.1177/000486580103400103.

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences, 25*, 167-177. https://www.sciencedirect.com/science/article/abs/pii/S0191886998000014?via%3Dihub.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, 59 (20), 22–57. https://pubmed.ncbi.nlm.nih.gov/9881538/.

Slade, M. (2010). Mental illness and well-being: the central importance of positive psychology and recovery approaches. *BMC health services research*, *10*, 26. https://doi.org/10.1186/1472-6963-10-26.

Smeijsters, H., & Cleven, G. (2006). The treatment of aggression using arts therapies in forensic psychiatry: Results of a qualitative inquiry. *The Arts in Psychotherapy, 33*. 37-58. 10.1016/j.aip.2005.07.001.

Smeijsters, H., Kil, J., Kurstjens, H. Welten, J., & Willemars, G. (2011). Arts Therapies For Young Offenders In Secure Care – A Practice-Based Research. *The Arts In Psychotherapy,* 38 (1), 41-51. https://psycnet.apa.org/record/2011-03161-006.

Smith, J.A., & Osborn, M. (2003). Interpretative phenomenological analysis. In J.A. Smith (Ed.), *Qualitative psychology: A practical guide to methods* (pp. 53-80). London, UK: Sage

Stahler, W. (2007). Prayerformance: A drama therapy approach with female prisoners recovering from addiction. *Journal of Creativity in Mental Health,* 2 (1), 3-12. https://www.tandfonline.com/doi/abs/10.1300/J456v02n01_02.

Stallone, T. M. (1993). The effects of psychodrama on inmates within a structured residential behavior modification program. *Journal of Group Psychotherapy, Psychodrama and Sociometry, 46*, 24-31. https://psycnet.apa.org/record/1994-03347-001.

Steinert, C., Hofmann, M., Kruse, J., & Leichsenring, F. (2014). Relapse Rates After Psychotherapy For Depression – Stable Long-Term Effects? A Meta-Analysis. *Journal Of Affective Disorders,* 168, 107-118. https://pubmed.ncbi.nlm.nih.gov/25043322/.

Stephenson, Z., & Watson, A. (2018). Scratching The Surface: A Service Evaluation Of An Applied Theatre Intervention For Female Offenders. *Prison Service Journal: Special Edition – The Arts In Prison,* 239, 16 – 21. https://www.crimeandjustice.org.uk/publications/psj/prison-service-journal-239.

Stewart-Brown, S., & Janmohamed, K. (2008). *Warwick-Edinburgh Mental Well-Being Scale (Wemwbs) User Guide, Version 1.* Http://Www.Healthscotland.Com/Documents/2702.Aspx.

Stewart-Brown, S., Janmohamed, K., & Parkinson, J. (2008). Warwick and Edinburgh Mental Well-being Scale (WEMWBS): User Guide, Version 1. NHS Health Scotland: Glasgow. https://hqlo.biomedcentral.com/articles/10.1186/1477-7525-5-63.

Stewart-Brown, S., Tennant, A., Tennant, R., Platt, S., Parkinson, J., & Weich, S. (2009). Internal Construct Validity Of The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): A Rasch Analysis Using Data From The Scottish Health Education Population Survey. *Health Quality Life Outcomes*, 7, 15. https://hqlo.biomedcentral.com/articles/10.1186/1477-7525-7-15.

Storbeck, J., & Clore, G. L. (2007). On the interdependence of cognition and emotion.

Streiner, D.L. & Norman, G.R. (2008). *Health Measurement Scales: A Practical Guide To Their Development And Use.* (4th Edition). Oxford University Press, Oxford. https://oxfordmedicine.com/view/10.1093/med/9780199685219.001.0001/med-9780199685219.

Taber, K.S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research Science Education. 48*, 1273–1296. https://doi.org/10.1007/s11165-016-9602-2.

Taggart, F. M, Friede, T., Weich S., Clarke, A., Johnson, M., & Stewart-Brown, S. (2013). Cross Cultural Evaluation Of The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) – A Mixed Methods Study. *Health Quality Life Outcomes, 11,* 27. https://hqlo.biomedcentral.com/articles/10.1186/1477-7525-11-27.

Teasdale, J. D. (1997). The relationship between cognition and emotion: The mind-in-place in mood disorders. In D. M. Clark & C. G. Fairburn (Eds.), *Oxford Medical Publications. Science And Practice Of Cognitive Behaviour Therapy* (p. 67–93). Oxford University Press.

Tennant, R., Fishwick, R., Platt, S., Joseph, S., & Stewart-Brown, S. (2006). *Monitoring Positive Mental Health In Scotland: Validating The Affectometer 2 Scale And Developing The Warwick-Edinburgh Mental Well-Being Scale For The UK.* Edinburgh, NHS Health Scotland. https://link.springer.com/article/10.1186/1477-7525-5-63.

Tennant, R., Hiller L., Fishwick R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007). The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): Development And UK Validation. *Health Quality Life Outcomes, 5*, 63. https://link.springer.com/article/10.1186/1477-7525-5-63.

The British Psychodrama Association (2020). https://www.psychodrama.org.uk/about/.

Thompson, J. (2002). *Drama workshops for anger management and offending behaviour*, London: Jessica Kingsley Publishers.

Tomyn, A., Fuller-Tyszkiewicz, M., & Cummins, R. (2011). The Personal Wellbeing Index: Psychometric Equivalence for Adults and School Children. *Social Indicators Research,* 110 (3), 913-924. https://www.researchgate.net/publication/256066108_The_personal_wellbeing_index_psychometric_equivalance_for_adults_and_school_children.

Volker, F., & Galbraith, N. (2018). Psychological therapy in prisons: Professionals' perceptions, *European Journal of Psychotherapy & Counselling, 20* (3), 278-293. 10.1080/13642537.2018.1495243.

Ward, T., & Brown, M., (2004). The Good Lives Model and conceptual issues in offender rehabilitation. Psychology Crime & Law, 10, 243-257. https://www.tandfonline.com/doi/abs/10.1080/10683160410001662744.

Ward, T., Day, A., Howells, K., & Birgden, A. (2004). The multifactor offender readiness model. *Aggression and Violent Behaviour, 9*, 645-67.

Watson, A. (2009). 'Lift your mask': Geese Theatre Company in performance. In T. Prentki & S. Preston (Eds.) *The Applied Theatre Reader* (pp. 47 – 54). Oxon, UK: Routledge.

Watson, D., Clark, L.A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 6 (*54), 1063 - 1070. https://psycnet.apa.org/doiLanding?doi=10.1037%2F0022-3514.54.6.1063.

Wetzel, E, Böhnke, J., & Brown, A. (2016). Response biases. https://www.researchgate.net/publication/282667681_Response_biases

Willis, G., Prescott, D., & Yates, P.M. (2013). *The Good Lives Model (GLM) in theory and practice. Computer Science.*

APPENDICES

Appendix A

Quality tool for mixed/quantitative studies

Question	Yes	No	Somewhat	Unsure. Any comments
Introduction				
Were the aims/objectives of the study clearly reported?				
Methods				
Was the study design appropriate for the reported aim(s)?				
Was the sample size justified for the study?				
Was the target population clearly defined?				
Was the sample frame taken from an appropriate population representative of the target population being investigated?				
Was the selection process likely to select participants that were representative of the target population being investigated?				
Were measures undertaken to address and categorise non-responders?				
Were the risk factor and outcome variables measured appropriate to the aims of the study?				
Were the risk factors and outcome variables measured correctly using established tools?				
Is it clear what was used to determine statistical significance (p-values, confidence intervals)				
Were the methods (including statistical methods) sufficiently described to enable them to be repeated?				
Results section				
Was the baseline data adequately described?				
Does the response rate raise concerns about non-response bias?				

If appropriate, was information about non-responders described?				
Were the results internally consistent?				
Were the results presented for all the analyses described in the methodology section?				
Discussion section				
Were the researchers' discussions and conclusions justified by the findings?				
Were the limitations of the study reported?				
Other				
Were there any conflicts of interest that may affect the authors' interpretation of the results?				
Was ethical approval or consent of participants obtained?				

Quality score: /20. No of 'unsure' items:

Adapted from: Downes, M. J., Brennan, M. L., Williams, H. C., & Dean, R. S. (2016). Development Of A Critical Appraisal Tool To Assess The Quality Of Cross-Sectional Studies (AXIS). BMJ Open, 6(12), 1-7. Doi:10.1136/Bmjopen-2016-011458.

Appendix B Quality tool for qualitative studies

Question	Yes	No	Somewhat	Unsure. Any comments
Introduction				
Was there a clear statement of aims/objectives of the study?				
Methods				
Was the qualitative methodology appropriate?				
Was the research design appropriate to address the aims of the research?				
Was the recruitment strategy appropriate for the aims of the study?				
Was the data collected in a way that addressed the research question?				
Has the relationship between the researcher and participants been considered?				
Results section				
Are the results valid?				
Was the data analysis sufficiently robust?				
Discussion section				
Was there a clear statement of findings?				
How valuable was the research?				
Other				
Have ethical issues been considered?				
Did the researcher consider their own potential impact on data collection, data analysis, and their findings?				

Quality score: / 11. No of 'unsure' items:

Adapted From: Critical Appraisal Skills Programme. (2017). CASP Qualitative Research Checklist. Retrieved From Http://Docs.Wixstatic.Com/Ugd/Dded87_25658615020e427da194a325e7773d42.Pdf.

Appendix C: Quality assessment of all included studies

Study Author		Quality criteria score	Quantitative/Qualitative/Mixed Methods Design
1	Blacker, Watson, & Beech, 2008	18/20	Quantitative
2	Colquhoun, Lord & Bacon, 2018	8/11	Qualitative
3	Gordon, Shenar, & Pendzik, 2018	8/11	Qualitative
4	Harkins, Pritchard, Haskayne, Watson, & Beech, 2011	17/20 9/11	Quantitative And Qualitative
5	Keulen-De Vos, Van Den Broek, Bernstein, Vallentin, & Arntz, 2017	15/20	Quantitative
6	Koiv & Kaudne, 2015	17/20 9/11	Quantitative And Qualitative
7	Reiss, Quayle, Brett & Meux, 1998	18/20	Quantitative
8	Mundt, Marín, Gabrysch, Sepúlveda, Roumeau, & Heritage, 2019	17/20	Quantitative
9	Stahler, 2007	19/20 9/11	Quantitative And Qualitative
10	Stephenson & Watson, 2018	19/20 10/11	Quantitative And Qualitative

Appendix D Data extraction tool

Study title	
Year of publication	
Author	
Country study took place in	
Characteristics of study	
Study Design/Aims	
Measure(S) Used/Data Collection Method(S) Used	
Measures: Standardisation; Validity; Reliability	
Participant Characteristics	
Sample size	
Gender of participants	
Age range and mean	
Ethnicity	
Response rate	
Method of recruitment	
Study findings	
Analysis used	
Methods used	
Findings and significance (quantitative)	

Themes appropriate to the aims/objectives questions (qualitative)	
Further relevant comments/information	
Review conclusion/Decision	

Adapted From: Fleeman, N., & Dundar, Y. (2014). Data Extraction: Where Do I Begin? In A. Boland, M. G. Cherry, & R. Dickson (Eds.), Doing A Systematic Review: A Student's Guide (Pp. 85- 98). London, UK: Sage.

Appendix E

Consent Form

Project Title: The Geese Theatre Company: An exploration into Actor Practitioners' (AP's) views and experiences of the use of Applied Theatre with offenders

Please read the following, and initial to confirm:

☐ I confirm I have read and understood the 'Participant Information Sheet' for the study. I have been given time to review the information and to ask the researcher questions.

☐ I understand that I am participating voluntarily, and I am able to withdraw from the research (within the given time period of one week) if I wish to do so without having to give a reason.

☐ I understand that the interview will be audio-recorded so the researcher has a clear account of what has been said. I also understand that only the researcher will have access to the recording and that the recording will be deleted as soon as it has been transcribed.

☐ I understand that the interview will be transcribed, and only the researcher and her supervisor will have access to the transcripts for the purposes of analyses.

☐ I have been informed of how my data will be stored in the 'Participant Information Sheet'.

☐ I understand that information I give during my interview will be used in a research report being submitted to the University of Birmingham which forms part of the researcher's doctorate degree. The research report will be made available to parties noted in the 'Participant Information Sheet', and could be published in an academic journal at some point in the future.

☐ It has been explained to me that my name will not be included in any write-ups of this project, and at no point will I be personally identified.

☐ I agree to take part in the above study.

Participant:

Date:

Signature:

Researcher:

Date:

Signature:

Appendix F

Introductory Letter to Potential Participants

I am conducting a piece of research exploring past and present actor practitioners' views, perspectives, and experiences of the use of applied theatre in forensic settings and would like to invite you to take part.

I have attached a 'Participant Information Sheet' which provides details of the project to help you decide if you'd like to take part.

As a Forensic Psychologist whom has worked therapeutically with forensic populations, I have a genuine interest in gaining an understanding of your views and experiences and feel this will benefit others.

My hope for this research is to give you a platform to discuss the important work you do. The wider aim of this project is to have a positive impact on those in forensic settings.

I would really appreciate it if you would kindly take a few minutes to read about the project, and please contact me on the below e-mail address if you would be happy to take part in a phone interview with myself at a time convenient for you. Also feel free to e-mail me or my supervisor with any questions about the project.

Many thanks, and kind regards,

Lucy Brierley, Forensic Psychologist

Researcher contact details: E-mail: [redacted] Tel: [redacted]

Supervisor contact details: Dr Zoe Stephenson
Lecturer in Forensic Psychology University of Birmingham

E-mail: [redacted] Tel: [redacted]

Appendix G

Demographic Information Sheet

Project Title: The Geese Theatre Company: An exploration into Actor Practitioners' (AP's) views, perspectives, and experiences of the use of Applied Theatre with offenders

How long have you been an AP?

If you are an ex member of staff, how long were you an AP?

How long ago did you leave your position of an AP?

Appendix H

Participant Debrief Sheet

Project Title: The Geese Theatre Company: An exploration into Actor Practitioners' (AP's) views, perspectives, and experiences of the use of Applied Theatre with offenders

I would like to take this opportunity to thank you for letting me interview you for my project. Your contributions will help further the understanding of the valued role that you as an Actor Practitioner (AP) hold.

The aim of this project is to explore the views, perspective and experiences of AP's in their role. Of particular interest is AP's experience of alliance building, i.e. how they build bonds, how they are perceived (as someone trustworthy), differences in interpersonal style (re actor practitioners' vs trained therapists), motivation to engage, etc...

Data from interviews will be analysed, where I will be looking for developing themes. The findings will be presented in a report for my doctorate degree.

The findings from this project will also help the directors of GTC understand the experiences of their AP's when undertaking this role.

If you would like to withdraw your data from this project, please let me know within one week of today. As such, after this date your interview will have been included in the final analysis.

If you choose to withdraw your data from this project there will be no negative consequences for you. If you have any concerns that I (the researcher) have been unable to address you are very welcome to contact my supervisor (details below). Similarly, if the interview has raised any difficult issues for you, you can discuss this with myself now or contact my project supervisor: Dr Zoe Stephenson at your convenience.

Please contact me should you wish to receive a copy of the final report (details below). Lastly, thank you once again for taking the time to participate in this project.

Many thanks, Lucy Brierley

Researcher contact details: Lucy Brierley, Forensic Psychologist

E-mail: [redacted] Tel: [redacted]

Supervisor contact details: Dr Zoe Stephenson, Lecturer in Forensic Psychology University of Birmingham

E-mail: [redacted] Tel: [redacted]

Appendix I

Semi-Structured Interview Question Themes

Project Title:

The Geese Theatre Company: An exploration into Actor Practitioners' (AP's) views and experiences of the use of Applied Theatre with offenders

In order to address the research question, a semi structured interview schedule has been developed which will incorporate the following:

How many years exp do you have? How long ago did you leave GTC?

1. How did you come to be in your role? (prompts: what did you do before this? Why did you apply?)
2. What are your thoughts on the training you had to do for the job? (Prompts: how long? Did you feel it prepared you for your role?)
3. How do you go about building a rapport with participants? My thought - Need a prompt
4. Can you tell me your thoughts on elements of the intervention you believe to be more successful to the participants?
5. Can you tell me about the types of interpersonal skills required from participants, in your opinion, needed/required for successful engagement with the intervention?
6. Have you had any particularly memorable experiences/interactions you during your role that you would like to talk about? – people took this of pos in the main
7. What's the appeal of working with people in prison / working with people with criminal justice histories?
8. What are some of the challenges about working with GTC? What are some of the rewards?
9. What are the range of potential benefits that participants might gain as result of attending a Geese programme? What are some of the potential outcomes for individuals?
10. What can an arts / creative approach provide that other interventions might not? – outside of the arts….

Appendix J

Participant Information Sheet

Project Title: The Geese Theatre Company: An exploration into Actor Practitioners' (AP's) views, perspectives, and experiences of the use of Applied Theatre with offenders

What is the purpose of this project?

The aim of this project is to explore the views, perspectives, and experiences of both current and past Geese Theatre Company Actor Practitioners (AP's) regarding the use of applied theatre programmes with offending populations.

Why have I been asked to take part?

I have asked you to take part in this project as you are either an active or ex AP with experience of using applied theatre in forensic settings. As such, I would like to interview you about your experiences in this role.

What will I be asked to do if I take part?

I will ask you to take part in an interview with myself (the researcher – Lucy Brierley) about your experiences of being an AP. The interview will be conducted via telephone and will be arranged at a time that will be convenient for you.

The interview will take approximately 45 minutes – 1 -hour, however depending on your responses, it may be a little shorter or longer.

Before the interview, I will ask you to read and sign a consent form to confirm that you agree to take part in the project. Once consent has been obtained, I will audio-record the interview using a dictaphone. The reason for this is so I am able to thoroughly and accurately capture everything you say.

Do I have to take part?

You absolutely do not have to take part in this project. It is solely your choice whether you take part. There will be no negative consequences of not taking part.

If you do take part, you are free to change your mind and withdraw from the project at any time during the interview. You are further free to withdraw at any time up to one-week following the interview. Following this time, your data from the interview would have been included in the data analysis so you will not be able to withdraw.

If you change your mind, no reason why will need to be provided. I will ensure that any information you have provided will not be used in the project. Your audio- recording, along with paper documents will be destroyed.

What are the benefits of my taking part?

There has been a lot of research into the effectiveness of GT programmes with offenders (e.g. reduction in aggression, depression etc…). However, there is currently no published review of the views and perspectives of AP's delivering these programmes. Although there is no direct benefit to you, we feel that other practitioners who work in forensic settings in various roles, could benefit from an insight into your expertise in working with offenders. Therefore, we feel that your participation in the project will ultimately benefit individuals within the criminal justice system.

What are the disadvantages or risks of my taking part?

We think it is very unlikely that there would be any disadvantages or risks to taking part. If there were any questions triggered any distress, then you do not have to answer. At the end of the interview, you will be debriefed by myself, and will have the opportunity to ask any questions/make any further comments.

Will my taking part in the study be kept confidential?

Under no circumstances will your name/personal details be revealed in any report generated from this project. Your name will appear on the consent form you sign only, and names will then be coded i.e. not identifiable. I alone will hold a record of participant names with their assigned code noted next to it. All other information (written and audio) will only have your assigned code attached to it.

Myself, along with my supervisor, Dr Zoe Stephenson at the University of Birmingham, will be the only ones with access to the information you provide. The information will be kept confidential and secure as detailed below:

Interview audio-recordings: will be saved to the University of Birmingham's secure IT system and immediately deleted from the Dictaphone. The interview transcript will be typed up and anonymised.

Information held electronically: (list of participant names, assigned codes, and interview transcript) will be saved on the University of Birmingham's secure IT system.

Written information: Your consent form will be printed and stored in a locked filing cabinet in the research supervisor's locked office. It will be deleted from e-mail and the computer. Only the researcher and the researcher's supervisor will have access to this locked cabinet.

The demographic information sheet will be destroyed following the information being saved onto the University of Birmingham's secure IT system. This will take place within the first 48 hours of the interview being conducted. This information will be summarised in the write-up but will not be attached to any specific quote to ensure that no individual could be identified by, for example, the length of time they have worked at GTC.

Quotes from your interview may be used in the write up but no information by which you could be identified would be included.

In line with the University of Birmingham's policy, the information you provide will be securely stored for 10 years. After this time, it will be destroyed by shredding and deleting.

What happens following the research?

A report will be written after I have transcribed all of the interview data. I will be looking for common themes reported by the interviewees about views and experiences of being an AP delivering drama and theatre based projects in criminal justice settings. My final report will be part of my thesis that I will be submitting for a doctorate degree at the University of Birmingham. The report will further be submitted to the GTC. It is hoped that the findings from this research may be published in a journal. As highlighted above, you will not be identified as a participant in any of these reports.

End note

This project is entirely my work and will contribute towards my doctorate degree at the University of Birmingham. I am working under the supervision of Dr Zoe Stephenson at the University of Birmingham.

Many thanks for taking the time to read about this project.

You are welcome to contact me should you have any questions.

If you would like to take part, can I request that you please contact as soon as you can.

Many thanks, Lucy Brierley

Researcher contact details:

Lucy Brierley, Forensic Psychologist

e-mail: [redacted] Tel: [redacted]

Supervisor contact details:

Dr Zoe Stephenson
Lecturer in Forensic Psychology University of Birmingham

E-mail [redacted] Tel: [redacted]

Appendix K Search syntax

prison AND drama*
prison AND theatre*
prison AND creative*
prison AND art*
prison AND psychodrama*
offender AND drama*
offender AND theatre*
offender AND creative*
offender AND art*
offender AND psychodrama*
inmate AND drama*
inmate AND theatre*
inmate AND creative*
inmate AND art*
inmate AND psychodrama*
criminal AND drama*
criminal AND theatre*
criminal AND creative*
criminal AND art*
criminal AND psychodrama*

Appendix L Personal Reflections on the Study Process

I was concerned that there may have been memory issues if the former APs hadn't worked for GTC for some time. However, when I conducted the interviews my worries were alleviated as the amount of information given by all APs was vast. That said, I was still mindful in the write-up of results that the time lapse between working for GTC and taking part in the interview may have had an impact on the information they gave. All interviewed gave fantastic examples of their work, and all explained how valuable their time at GTC had been – there was a very positive feel to the interviews. Upon further reflection, I considered that former APs may have actually felt more able to disclose any negative issues that they may have experienced as opposed to current APs who may have been concerned about any negative consequences of painting their organisation in a bad light. I felt that they may have been more objective about the work than current APs, however, this is unknown, and I was not working on the assumption that current APs would consciously deceive me regarding their work but merely that there may have been some potential for bias. Given that former and current APs reported similar views and experiences, I felt more confident that these biases, if any, were small. The only negative comments were around having to be away from family for a long time as they described a travelling theatre company on the road the for sometimes months at a time.

All participants spoke about the positive aspects of the interventions rather than focusing/reporting cases where participants did not engage or dropped out of the intervention. On reflection, this may be as they were hoping to highlight the positive impact their work has given the amount of effort and expertise they put into their roles. Thinking about instances where the intervention did not 'work' would perhaps cause some distress. That said, APs did acknowledge that they experienced challenges and people who did not want to engage so it wasn't felt that they were completely one-sided in their views. Additionally, it may well be the case (as is supported by literature) that the interventions do have a positive impact on many

people and, as such, it is understandable that many positive examples are given. I was clear that I was taking a more semantic/realist approach to the data so, to an extent, accepted what participants said as being true. However, I was aware that, as with any qualitative study, there will be societal/psychological factors that may have impacted responses.

A further concern was that, given my lack of experience in conducting research such as this, I would struggle to get a lot of information from participants; that they wouldn't be forthcoming with information. However, in all interviews I found everyone to be very personable and chatty which put me at ease. On reflection this could be in part due to their roles as actors, i.e., perhaps actors are more likely to be forthcoming and confident.

Interviewing process: my potential biases

Whilst I am aware that research shows more favourable towards face to face interviewing, this was not an option for this study due to the accessibility of the APs (i.e., some of them were not residing in the UK and there was a need to socially distance due to C-19). It was also the case that current APs were furloughed due to C-19. I did consider using skype/zoom as a medium to interview participants, however I was concerned that signal issues may have arisen especially with overseas participants, and a jumpy signal I feared may compromise momentum. I was concerned about building rapport with participants on the telephone, as with no social cues or nonverbal nuisances as noted in papers such as Novick's (2008) may have compromised the quality of the data. However, this did not seem to be an issue. I also felt that the telephone medium can be less formal and may put people at ease in disclosing information. I therefore do not think that this method had an adverse impact on the quality of the data I obtained.

I was mindful that I was coming to the interviews and analysis with my own biases. As a forensic psychologist and being aware of the literature on DbT, I was aware that I have a

positive view of DbT rather than being completely objective about it. I was also aware that the GTC have worked very hard to build up their excellent reputation in the UK and I would not have wanted to report anything negative/concerning about the company. I ensured when designing the interview questions that participants weren't asked leading questions and was mindful during interviews that I should allow them to do the vast majority of the talking – thus hopefully minimising any influence I might have on them. During data analysis I prepared myself for the potential need to write less than positive information, however, information provided by participants about the GTC was purely positive, so this did not become an issue.

Thematic Analysis process

When I reflect on the process of coding and grouping the data, I was proud of the amount of data I obtained, and the quality of the data. I was especially excited by the number of examples of the impact of the great work APs do/did. The process of the development of themes and subthemes was a long one. Much of the data overlapped and so it was difficult applying cut offs to similar data (i.e. making decisions about which theme/sub-theme information would be placed). I was concerned at the beginning of this process that I wouldn't be able to tease data apart from each other to see themes emerge. However, from following the TA analysis guidance, spending a lot of time immersing myself in the data, and discussion with my academic supervisor, I was able to develop clear and defined themes and subthemes. In terms of how many interviews were an acceptable number to hold integrity of the project, this was less clear. I would have liked to have at least twelve interviews but because of C-19 I only had seven participants. Although this was disappointing, I was relieved that it was deemed acceptable by the UoB research team. In addition, I consulted the literature to further understand data power. In the new TA reflexive literature, the guidance outlines that a researcher should be guided by the quality of each data set and that although this is subjective

from one researcher to the next, a decision cannot be made until the data is generated and studied for content, context and integrity (Braun & Clarke, 2019). The quality of the data seemed high/detailed. As such, although I would still like to include more participants in the study, I felt that there was enough data to generate some interesting/useful findings.