Assessing the fidelity of delivery style of a mental skills training programme for young people experiencing homelessness

Tidmarsh, Grace; Whiting, Richard; Thompson, Janice L; Cumming, Jennifer

DOI:
10.1016/j.evalprogplan.2022.102150

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Tidmarsh, G, Whiting, R, Thompson, JL & Cumming, J 2022, 'Assessing the fidelity of delivery style of a mental skills training programme for young people experiencing homelessness', *Evaluation and Program Planning*, vol. 94, 102150. https://doi.org/10.1016/j.evalprogplan.2022.102150

Evaluation and Program Planning 94 (2022) 102150

Contents lists available at ScienceDirect

Evaluation and Program Planning

journal homepage: www.elsevier.com/locate/evalprogplan

Assessing the fidelity of delivery style of a mental skills training programme for young people experiencing homelessness

Grace Tidmarsh [a,*,1], Richard Whiting [a,2], Janice L. Thompson [a,3], Jennifer Cumming [a,b,4]

[a] School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, United Kingdom
[b] Institute for Mental health, University of Birmingham, United Kingdom

ARTICLE INFO

Keywords:
Homelessness
Youth
Positive Youth Development
Mixed Methods and Process Evaluation

ABSTRACT

There is a need for positive youth development/strengths-based approaches to support the wellbeing and social inclusion of young people experiencing or at risk of homelessness. My Strengths Training for Life™ (MST4Life™) uses a strengths-based approach with the aim to improve young people's resilience, self-worth, wellbeing and engagement in education, employment, and training. This mixed methods study assessed the fidelity of delivery style of the MST4Life™ programme, the extent to which frontline service staff can delivery psychologically informed programmes to service users and identified barriers and enablers to delivering with fidelity to the intended style. Observations of programme delivery (two facilitators per session) took place across early, middle, and late phases of the programme across a pilot phase (n = 18) and main study (n = 45). Facilitators also completed self-reflection forms following each session. The mean observation score was 82.2 ± 15.7 %, and facilitator self-report mean adherence score was 89.3 ± 6.2 % which indicate that the programme was delivered with high fidelity. Quantitative data was also analysed using non-parametric statistical test (Mann-Whitney U Test). There was a significant difference between observation scores for deliverers with postgraduate psychology training compared to deliverers without postgraduate psychology training (p = .029). Qualitative data were analysed using inductive thematic analysis. Barriers and enablers included communication, frontline staff support, logistics, and participant behaviours. Overall, this study highlights that despite the challenges of delivering complex community programmes to young people experiencing homelessness, it was possible for frontline service staff to deliver MST4Life™ with high fidelity.

1. Introduction

Many interventions utilise outcome evaluations to demonstrate effectiveness in achieving programme goals (Anderson et al., 2013; Pettee Gabriel, DiGioacchino DeBate, High, & Racine, 2011). Conclusions around intervention effectiveness are often based upon outcome results and procedures described in the methods, rather than on procedures implemented (Dobson & Cook, 1980). The assumption that an intervention has been delivered as described can present challenges for researchers and practitioners during implementation. These include making content changes due to time constraints, adaptations to meet participant needs, and/or insufficient training for deliverers. Less commonly done, process evaluations are critical to understanding how programmes are implemented and whether implementation challenges may account for the variability in programme impact (Iachini, Beets, Ball, & Lohman, 2014).

Encompassed within process evaluations are fidelity assessments, which are vital to measuring adherence to programme implementation and enabling outcomes to be correctly attributed (or not) to interventions. Evaluators use fidelity assessments to: (a) ensure the intervention is delivered in line with the protocol; (b) enable a more thorough understanding of effectiveness of complex interventions; (c) provide programme developers with the details of what is delivered within the programme sessions; and (d) create a platform for potential

* Corresponding author at: School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, B15 2TT, United Kingdom.
E-mail address: G.tidmarsh@bham.ac.uk (G. Tidmarsh).
[1] ORCID ID: 0000-0001-8833-1960.
[2] ORCID ID: 0000-0003-4045-4376.
[3] ORCID ID: 0000-0002-7312-3226.
[4] ORCID ID: 0000-0001-5655-7842.

https://doi.org/10.1016/j.evalprogplan.2022.102150
Received 13 February 2021; Received in revised form 10 March 2022; Accepted 1 August 2022
Available online 3 August 2022

improvements post-intervention and for real-time corrections during intervention delivery (Kaye & Osteen, 2011; Walton, Spector, Tombor, & Michie, 2017). Fidelity assessments provide a platform for researchers and stakeholders to gain an in-depth understanding of programme implementation and make evidence-based alterations to delivery where needed. Experts have highlighted that fidelity assessments are rarely implemented due to their cost and time-consuming nature (Hardeman et al., 2008; Borrelli, 2011). But, without these assessments, there is a risk of ineffective interventions informing policy or clinical practice, resulting in higher economic, societal, and scientific costs (Bellg, 2004; Borrelli, 2011). Therefore, experts advocate the use of fidelity assessments to inform policy and practice and inform more effective use of funding (Bruns, Burchard, Suter, Leverentz-Brady, & Force, 2004).

1.1. Methods of implementing fidelity assessments

Fidelity can be part of a full process evaluation or conducted as a standalone piece of research to assess one or more of the following: quality of delivery (e.g., adherence to session content and adherence to intended delivery style), dosage of delivery received, participant responsiveness (engagement), recruitment, and intervention complexity (Borrelli, 2011; Hardeman et al., 2008; Hasson, 2010; Mihalic, Fagan, & Argamaso, 2008). Used in isolation or in combination to improve rigour, evaluators can employ a variety of methodological tools including audio-recording, video-recording, self-report, real-time observations, rating scales, and questionnaires. The use of real-time observations, audio-recording and self-report have been applied successfully in educational (Maynard, Peters, Vaughn, & Sarteschi, 2013) and health settings (Rixon et al., 2016). Employing multiple methods also provides evaluators with an opportunity to draw comparisons. For instance, Hardemen et al. (2008) audio-recorded sessions in a physical activity behaviour change intervention to assess adherence to delivery behaviours and collected self-report data from facilitators. Combining these methods enabled the researchers to compare actual delivery with perceived delivery.

Evaluators must also consider the frequency, time points (e.g., beginning, middle, and end of programme), and duration of data collection (e.g., entire or part). Collecting data throughout the intervention ensures data are representative of the entire programme, as is reporting sampling and analysis techniques in full (Kitzman-Ulrich et al., 2009). In their evaluation of the "Active by Choice Today" randomised school-based trial, Kitzman-Ulrich et al. (2009) conducted observations in 2-week blocks (3 sessions/week) during Weeks 1 and 2, 8 and 9, and 15 and 16 of a 17-week programme. The staggered start at each location enabled observations to be conducted at different sites. Thus, evaluators observed the same time points for each programme as well as collected data across the full programme, which enhanced the understanding of programme mechanisms.

1.2. Challenges of delivering complex interventions with fidelity

Fidelity assessments are especially important when evaluating complex interventions where multiple factors can influence the extent to which fidelity of delivery style is achieved. A complex intervention comprises of multiple interacting components, although additional dimensions of complexity could include the implementation difficulty and the number of organisational levels targeted (Moore et al., 2015). Due to their nature, complex interventions usually undergo some degree of tailoring when implemented in new contexts. Capturing what is delivered in practice, with close reference to intervention theory, can enable evaluators to distinguish between adaptations made to the intervention to fit different contexts as compared to changes that undermine intervention fidelity (Bumbarger & Perkins, 2008; Hawe, Shiell, Riley, & Gold, 2004; Moore et al., 2015).

Tailoring of complex interventions is important for meeting the varying needs of participants. However, tailoring complex interventions can present facilitators with several challenges to delivering with fidelity. Capturing these challenges enables discussions around improving adherence to delivering the intervention as designed (Wierenga et al., 2013), yet remains an under-reported area of fidelity assessments. In a complex intervention for disadvantaged youth, Mihalic et al. (2008) described multiple barriers to delivering with fidelity including poor behaviour of participants, lack of classroom management skills from staff, and logistical challenges of delivering across multiple sites. Given the similarities of their target population to that of the present study, we expected to experience similar challenges in the delivery of a strengths-based intervention with young people experiencing homelessness, participating in the My Strengths Training for Life™ (MST4Life™) programme.

1.3. Fidelity of delivery style for interventions with young people experiencing homelessness

Homelessness among young people (16–24 years) is an increasing social and economic problem in the United Kingdom (UK). Frontline service providers are under pressure to support the rising number of young people experiencing homelessness whilst simultaneously experiencing cuts to their funding and a reduction in resources and capacity (Homeless Link, 2021). Furthermore, services have limited access to specialist mental health support and early intervention initiatives, resulting in a lack of ability to meet the complex needs of young people experiencing homelessness at a crucial time in their development towards independence and adult life (MacKie & Thomas, 2014). Strengths-based practice has been increasingly advocated over more traditional deficit-based/risk reduction approaches that emphasize negative outcomes and reinforce negative stereotypes (Cronley & Evans, 2017).

Although frontline service staff have experience in engaging with young people (YP) within housing services their roles have been predominantly focused on meeting YPs basic needs (e.g., housing, food, water, and crisis support) rather than providing more holistic support by incorporating long-term personal development and emotional needs. Moving to a roles that require staff to provide strengths-based holistic support requires shifts in peoples personal and professional frameworks, something which is not always easy (Blundo, 2001). Furthermore, in a study assessing the integration of strengths-based interventions in child welfare provision one challenge that was reported was that despite leaders being onboard staff did not always want to implement the strengths-based strategies (Sabalauskas, Ortolani, & McCall, 2014). To ensure that a strengths-based and psychologically informed approach is used in practice, assessing the fidelity of delivery style is key to understand if key concepts of the outlined approach are implemented. Within this study delivery style is defined as the intentional behaviours and actions of the people delivering (or supporting) a programme to create the social climate and atmosphere that enables participants to achieve positive outcomes. Within positive youth development (PYD), the climate has been recognised as the social environment that is created with and between adults, peers, and parents (Holt et al., 2017). However, few studies have investigated fidelity to delivery style, so it is not known what factors enable or inhibit facilitators to delivering a programme with high fidelity to the intended delivery style (Tidmarsh, Thompson, Quinton, & Cumming, 2022) and therefore the extent to which desired social climate is achieved.

Previous studies provide evidence of the benefits of continued support and training for staff (DuBois, Holloway, Valentine, & Cooper, 2002; McQuillin, Straight, & Saeki, 2015). Assessing the fidelity of delivery style acts as an ongoing feedback mechanism to identify programme facilitators' further training needs and if/when booster sessions are needed (Cumming et al., 2022). Ongoing support beyond initial training has been highlighted as a key strategy for ensuring fidelity (McQuillin et al., 2015). Studies within school-based youth mentorship programmes reported that ongoing support and training predicted

greater anticipation of continuing to mentor young people as well as producing stronger youth outcomes (DuBois et al., 2002; McQuillin & Lyons, 2021; McQuillin et al., 2015). As such ongoing support and training has the potential to contribute to overcoming barriers to delivering programmes with fidelity to delivery style reported in other PYD programmes for disadvantaged young people. These include challenges retaining staff, staff knowledge of programme components and style, and managing participant behaviour (Collins, Lavender, Brown, Sheffield, & Andrew Aligne, 2013; Kenyon et al., 2019; Mihalic et al., 2008). The extent to which evaluating the fidelity of delivery style through understanding barriers to delivery in strengths-based PYD programmes for young people experiencing homelessness is yet be explored but has the potential to further benefit programme facilitators in achieving delivery style expectations and therefore supporting positive youth outcomes.

Programs and approaches (including delivery style as defined in this paper) have the capacity to impede youth development as well as promote it (Roth & Brooks-Gunn, 2016). Encouraging autonomy in participants is achieved through a strengths-based delivery style created through the purposeful actions and behaviours of facilitators. In a study aimed at understanding organisations serving runaway and homeless youth in America, Gwadz et al. (2019) compared higher and lower quality services and reported that lower quality services focussed more on basic services and crisis support with less attention given to emotional support. In contrast, higher quality services focussed on short and long-term goals, developed youth centred environments, and encouraged autonomy. A key component of PYD programmes is that young people are approached as people who can develop rather than problems that need to be solved (Roth & Brooks-Gunn, 2003), a more recent view which has been adopted within the context of homelessness. Supporting young people in this manner is achieved through language and style, as much as the activities used within a PYD programme. As such understanding the extent to which delivery style is implemented is an essential component to creating high quality services that can address YPs complex needs beyond immediate crisis support, ensuring practices implemented promote not impede youth development.

With homelessness also having a detrimental impact on a young person's health, well-being, and future economic prospects (Homeless Link, 2021), there is a pressing need for effective interventions that can address the co-occurring challenges that these young people experience and be delivered within the constraints faced by public services. To address this gap, the MST4Life™ programme was co-produced with young people and staff from a homeless service in the UK to improve well-being and employability. It draws from aspects of sport psychology including mental skills training (MST), an approach more typically associated with elite sportspeople (Vealey, 2007), and is underpinned by strengths-based psychology through PYD (Krabbenborg, Boersma, & Wolf, 2013) and self-determination theory (SDT; Deci & Ryan, 1985; Deci, 2000). The programme is also framed in the housing services' psychologically informed environments (PIE) (Cumming, Skeate, & Anderson, 2017) framework.

2. Aims and hypotheses

The study aimed to explore the extent to which frontline service staff were able to deliver a psychologically informed intervention with high fidelity of delivery style. Due to the extensive psychological training and education possessed by programme deliverers compared to the co-deliverers previous experience predominantly in crisis support, it was hypothesised that the programme: (a) will be delivered with higher fidelity by programme deliverers with postgraduate level psychology training; and (b) will be delivered with lower fidelity by co-deliverers (frontline service staff). The objectives addressed to meet the study aim included: (a) comparing achieved fidelity scores from observations of the delivery team with their self-reported fidelity scores; and (b) qualitatively exploring the challenges and enablers to delivering a complex intervention with fidelity.

The present study set out to address the current gaps around conducting fidelity assessments of complex interventions for populations with high need support. Fidelity studies within the field of homelessness have focused on provision of housing at crisis point (Bernard, 2018; Rae et al., 2018). To our knowledge, this study is the first to assess the fidelity of delivery style within a programme delivered to young people experiencing homelessness, who are currently living in supported accommodation with the aim of improving their progression into independence. MST4Life™ was delivered by facilitators with postgraduate level training in psychology, and frontline service staff without in-depth training in psychology; thus, this evaluation specifically considered fidelity of delivery style. Understanding the extent to which it is possible for programmes such as MST4Life™ to be delivered with fidelity by frontline staff is key to understand the viability of frontline services adopting these programmes in the future.

3. Methods

3.1. Intervention setting

The current study was part of a larger ongoing outcome and process evaluation of MST4Life™ (see Parry, Quinton, Holland, Thompson, & Cumming, 2021; Quinton et al., 2021) an intervention implemented between 2014 and 2020 across 21 accommodation sites of the housing service. To ensure sustainability of MST4Life™, the delivery team expanded in the second half of the intervention to include front-line service staff and more robust evaluation (compared with only self-reflection forms and enjoyment ratings). The pilot and main fidelity assessment drew on 3 waves of programme delivery between April 2018 and June 2019. The pilot study aimed to develop the observation/self-reflection tool and adapt traditional observation processes for the main study. The housing service supports young people aged 16–25 who are experiencing homelessness or at risk of homelessness, helping over 5000 young people per year within the West Midlands, UK. The service has 39 supported accommodation schemes as well as a range of prevention, employability, and engagement services to help young people regain the stability needed to rebuild their lives, and gain skills, confidence, and employment.

3.2. Programme

MST4Life™ is for young people (aged 16–25 years) who are experiencing homelessness or at risk living in supported accommodation, and includes training for frontline staff (e.g., support workers, employability coaches) to deliver and support the programme. Involving staff in the design and delivery model is critical to supporting the long-term sustainability and impact of the intervention and improve engagement and uptake of service users to community based programmes such as MST4Life™ (Greenhalgh, Jackson, Shaw, & Janamian, 2016). Young people accessing MST4Life™ often have poor mental health and other high support needs (e.g., substance abuse, learning difficulties, pregnant, or young parent). Throughout the pilot phase of the fidelity assessment 34 young people were taking part in MST4Life™ and 53 during the main study. Total numbers of YP were taken from registers completed by facilitators. Demographic information was collected (see Table 1); however, it does not represent the total number of YP engaged in MST4Life™ during the fidelity assessment as not all YP completed questionnaires (38.2 % = pilot work, 54.7 % = main study).

MST4Life™ aims to provide participants with skills building opportunities through challenging and meaningful activities and build positive relationships with adults and peers (Cumming et al., 2022). It is a complex intervention consisting of two phases and delivered at multiple accommodations sites. Phase 1 involves 10 two-hour sessions in the young person's local service (sessions include individual and team tasks such as emotional regulation and air vehicle challenge). Phase 2 is a

Table 1
Table showing demographic information collected from YP engaged in MST4Life™ during the pilot and main study.

	Pilot Study	Main Study
Average Age ± SD (years)	19.0 ± 1.5	19.5 ± 2.5
Gender	Female = 10 Male = 3 Non-Binary = 1	Female = 7 Male = 13
Ethnicity	White = 4 Black/African/Caribbean/ Black British = 4 Other = 1	White = 18 Black/African/Caribbean/ Black British = 2 Mixed = 2 Arab = 1
EET/NEET Status*	EET = 6 NEET looking for work = 5 NEET not looking for work = 1	NEET looking for work = 1 NEET not looking for work = 1 Unable to work/other = 6

Note: Engaged in employment, education, or training = EET; Not engaged in employment, education, or training = NEET.

3-night/4-day outdoor adventure education residential (activities include canoeing, high ropes, and hiking). Delivery of phase 2 is assisted by trained outdoor instructors. In the short term, MST4Life™ aims to increase participants' mental skills, strengths and wellbeing and reduce engagement in risky behaviours. MST4Life™ is delivered by facilitators from an academic institution and staff from the homeless service trained as co-deliverers.

3.3. Study participants

3.3.1. Delivery team

The delivery team (N = 8) consisted of four deliverers (all with postgraduate psychology training) and four co-deliverers (frontline staff from the housing service, with undergraduate degrees in social care). All the delivery team received PIE training (Cumming et al., 2017) through the housing service as part of staff training and support. Deliverers (3 male, 1 female; M Age = 28.2 ± 4.8 years) had a wide range of experience working within MST (between 1 and 12 years) and had previously delivered at least 50 MST4Life™ sessions; more experienced facilitators had delivered over 200 sessions. Co-deliverers (1 male, 3 female; M age = 39.4 ± 9.6 years) had a range of experience working with youth (8–20 years) and had minimal previous exposure to MST4Life™ delivery sessions prior to this study (range = 5–10 sessions). Co-deliverers received three days of training from programme deliverers to develop their understanding of MST and its underpinning theories (SDT and PYD), delivery style, session content, and reflection techniques. Other housing service staff would also attend MST4Life™ sessions in the role of support workers for the YP, or to stay late where young people had requested MST4Life™ be delivered in the evening.

4. Procedures

A mixed methods approach assessed the fidelity of delivery style of MST4Life™ using systematic observation and self-reported ratings. The use of self-report measures is deemed an unreliable measure of fidelity when used alone due to over-reporting of fidelity (Breitenstein et al., 2010; Hardeman et al., 2008; Walton et al., 2017). However, self-report data can be useful to enable a holistic assessment of fidelity of delivery style. Facilitator perceptions of their adherence to delivery style are key to allowing the comparison to observed adherence, and to make effective amendments (Hardeman et al., 2008) and highlight where further training may be required. The same rating scale was used for both types of data to assess the degree to which programme facilitators delivered MST4Life™ in the intended style. Ethical approval was granted by the University ethics committee. Delivery team and observer consent were obtained prior to the study. YP engaging in MST4Life™ provided written consent. Verbal consent was obtained from the YP prior to the start of each observed session.

4.1. Observers

Three individuals (two female, one male, M Age = 24.5 ± 1.7 years) carried out face-to-face observations during the pilot work and main study. Face-to-face observations were considered most appropriate as nine YP participating in MST4Life™ did not consent to photographs or recordings being taken of them. The lead observer (female) carried out all observations during the pilot study. Additional observers were recruited and trained to conduct observations in the main study alongside the lead observer. Through previous work the lead observer had a good understanding of the unique challenges that can occur when working alongside YP with ill-mental health, complex needs, and traumatic past experiences; characteristics which were present within the recipients of the MST4Life™ programme. All observers had good knowledge of SDT (Ryan & Deci, 2000), possessing a sound understanding of the theory underpinning both the intervention and the observation tool.

After introducing themselves and checking all participants in the room were comfortable for the observation to proceed, the observer would move to the back of the room and observe the session with minimal interaction. Observations lasted for the entirety of each session, and a form was completed by the observer for each facilitator and co-deliverer present (n = 2 forms per session). Field notes were made throughout on behaviours and conduct that stood out as either good or poor practice. Direct quotes of facilitator's comments were also noted. Following the session, the rating scale was completed as well as a written overview of the session alongside any suggestions for improvements. Observations were not conducted in Sessions 1 or 2 of Phase 1 to allow the YP time to become familiar with each other and the facilitators before adding in another unknown factor.

4.2. Tool development

The observation and self-reflection scales were a bespoke tool developed to meet the needs of the MST4Life™ project and frontline housing service in which the programme is delivered.

The observation and self-reflection tools described above were developed using an iterative approach whereby alterations were made during and following the pilot study. Content validity was checked through working with university deliverers and frontline service staff through individual feedback and meetings to review and change the tool. A number of frontline staff noted the tool was quite long and that some items could have been more appropriately named. As a result, some items were removed where they were deemed not applicable to delivery style and also to ensure the tool remained a manageable size in terms of time to complete. For example, item 25 had initially been two separate items; behaviour management and group focus maintained and was combined to form one item.

The tool is underpinned by Deci and Ryan's Basic Needs Theory (BNT) (2002) and the final version comprised of 5 sections, totalling 27 items (see Table 2) to assess the extent to which facilitators displayed need supportive (e.g., providing opportunity for input and choice) or

Table 2
The Sections and items within the observation and self-report tool.

Section	Need supportive items	Need thwarting items	Total items for section
Competence	2	2	4
Autonomy	3	5	8
Relatedness	5	2	7
Communication	3	1	4
Structure	2	2	4
Entire tool	**15**	**12**	**27**

need thwarting behaviours (e.g., uses controlling language). Need support is where behaviours are exhibited that promote positive feelings of autonomy, competence, and relatedness, whereas thwarting is a behaviour (or behaviours) which directly inhibits the promotion of those needs or decreases current levels (Deci & Ryan, 2002). Each item was rated on a scale of 0–3 (0 = not at all demonstrated; 1 = Displayed, but in a limited way; 2 = Displayed with moderate frequency or conviction; 3 = consistently demonstrated with conviction). The maximum score possible for supportive behaviours is 45, whilst the maximum score for thwarting behaviours is 42. The total score for each participant was calculated by adding up the score from each item and then the total score was converted to a percentage. Thresholds for fidelity level achieved were set a-priori based on discussion amongst the research team and the housing service and were defined as follows for need supportive behaviours: low = ≤ 59 %, medium = 60–79 %, high = ≥ 80 %. These thresholds are similar to those seen elsewhere within fidelity assessments in an educational setting where programmes were delivered with adequate fidelity at 65 % (Lorentson, Joo Oh, & LaBanca, 2014) and 80 % (Balu & Quint, 2014).

Self-reflection forms were completed by facilitators and co-deliverers after each session (in the pilot and main study) during Phase 1, and an overall written reflection was provided following the outdoor adventure education (OAE) residential (Phase 2). Self-reflection forms were identical to the observation form. Deliverers also completed questions designed to facilitate written reflection of challenges and successes of the session, areas for improvement for future delivery, and any additional comments deliverers wished to make.

4.3. Data collection

4.3.1. Pilot study

A total of 18 observations were conducted across three accommodation sites and the OAE centre where MST4Life™ was delivered between April 2018 and June 2018. Sessions were purposively selected to represent early, middle, and late stages of programme delivery (Kitzman-Ulrich et al., 2009). In total, seven out of nine selected sessions were observed at accommodation sites. Two deliverers were observed per session. Similarly, three out of four sessions were observed at the OAE centre. Only one deliverer per session was observed at the OAE centre. Programme facilitators completed the rating tool following each session delivered in local accommodation (n = 47). Four were not completed due to session cancellation. The remaining self-reflections were not completed despite the session being delivered. In the pilot study, 30 % of sessions were systematically selected for observation. Cancellation of some sessions (n = 2) meant that the number of observed sessions was below the 30 % threshold. Thus, 40 % of sessions were selected for observation in the main study which is within the range recommended by Schlosser (2002).

4.3.2. Main study

The number of sessions delivered by each facilitator and co-deliverer and the number of sessions observed are reported in Table 3. In total 45 observations were completed between October 2018 and March 2019 across two programmes delivered over six accommodation sites.Two deliverers were observed during Phase 1 sessions and only one during Phase 2. Deliverers completed the rating tool following each session (n = 84).

5. Data analyses

Data from observations, self-report questionnaires, and written reflections from the main study were analysed using both quantitative and qualitative methods. Quantitative data were analyzed using SPSS (Version 24, 2018). The non-parametric Mann-Whitney U test was used to analyse fidelity scores, as these data were not normally distributed (i.e., data were nominal or ordinal). Qualitative data were extracted from the observation and self-report tools and thematically analysed (Braun & Clarke, 2006, 2013). Familiarisation, generation of initial codes and initial themes were completed by the lead author using an inductive approach. Initial codes were created using NVIVO (Version 11, 2017) and developed further by hand. Rigour and trustworthiness were established through critical appraisal of themes to broaden the interpretation of the data beyond the first author. The second author and programme facilitators acted as critical friends who encouraged reflection on, and development of more nuanced reading of the data as well as challenging example quotes and themes (Smith & McGannon, 2018). The quantitative results (i.e., fidelity scores) are presented first, separating results from the pilot study and main study. Qualitative results are presented after the quantitative results; results from the pilot study and main study are combined.

Table 3
Data on the number of times facilitators delivered observations, and frequency of observations for each individual.

Project	Phase 1			Phase 2	
	Facilitator	Sessions facilitated	Number of times observed	Sessions facilitated	Number of times observed
Pilot	F1	15	4	6	1
	F2	16	5	6	1
	F3	16	4	6	1
	F4	6	2	N/A	N/A
Main study	F1	9	4	N/A	N/A
	F2	21	9	6	2
	F3	22	8	6	2
	F4	1	1	N/A	N/A
	CD1	13	5	N/A	N/A
	CD2	6	5	6	2
	CD3	13	5	N/A	N/A
	CD4	7	2	6	0

Note. N/A: Non-applicable for this facilitator as not present during Phase 2 of programme delivery.

6. Results

6.1. Fidelity scores

6.1.1. Pilot study

The mean overall adherence score of observed sessions was 89.9 % ± 8.8 and the overall adherence score of the facilitator self-report questionnaires was 96.6 % ± 3.9. Observations were on average 6 % lower than self-report scores.

6.1.2. Main study

In the main study scores from the observed sessions (N = 45) and the facilitator self-report questionnaires for the entire programme (N = 84) indicated high adherence to delivery style. The mean overall adherence score of observed sessions was 82.2 ± 15.7 %, and facilitator self-report mean adherence score was 89.3 ± 6.2 %. Observations were on average 9 % lower than the self-report scores, however both fidelity scores fall within the "high" category. Details of adherence scores are reported in Table 4.

Mann-Whitney U analyses indicated that observation scores of the more experienced university staff were significantly higher from frontline staff (p = .029). There was no significant difference between facilitators and frontline service staff for self-report scores (p = .114). Regarding level of psychology training, there was a significant difference between observation scores and the deliverers with postgraduate psychology training (p = .029), indicating that those with psychology backgrounds were able to deliver the programme with higher fidelity. There was no significant difference between self-report scores and level of psychology background (p = .114).

Table 4
Observation and self-report scores from each session.

Facilitator (employer)	Need Supportive Behaviours		Need Thwarting behaviours	
	Observed	Self-report	Observed	Self-report
F1	High (98 %)	High (92 %)	0 %	0.8 %
F2	High (97 %)	High (97 %)	0.2 %	0 %
F3	High (93 %)	High (93 %)	0.6 %	1 %
F4[a]	High (97 %)	–	0 %	–
CD1	Medium (71 %)	High (85 %)	0 %	5.5 %
CD2	Medium (69 %)	High (93 %)	3.7 %	7.5 %
CD3	Low (57 %)	High (86 %)	8 %	8.3 %
CD4	Medium (77 %)	Medium (79 %)	10.5 %	9.3 %
Combined average:	**High (82 %)**	**High (89 %)**	**3 %**	**4.3 %**

Note. F = facilitator from the University; CD = facilitator from the housing service acting as co-deliverer.

[a] Self-report data were not available for F4. Also, there are a number of self-report data missing for project 2, which was due to young people not attending, leading to cancellation of the session, in conjunction with F4 being a regular facilitator at this location.

6.2. Barriers and enablers influencing fidelity of delivery style

The main themes highlighting barriers to, and enablers of, fidelity of delivery style are reported in Table 5. The themes were inclusive of facilitator, participant, service provider staff, and outdoor instructors' behaviours, as well as factors outside of individuals' direct control, such as the weather. During the identification of the themes, it became apparent that some barriers and enablers were present only in Phase 1 and others only in Phase 2, whilst some were present across both phases (see Table 5).

6.2.1. Barriers to fidelity of delivery style

Poor communication, practical challenges, and participant behaviour were barriers to fidelity of delivery style that were only relevant during Phase 1 of MST4Life™. Facilitators perceived poor communication with YP as a barrier to delivery, such as attempting to explain a concept to a participant in a complex manner that was not understood by YP. Facilitators also discussed how poor communication between service provider staff and YP created a barrier to delivering in the desired style. For example, one facilitator discussed how staff had not encouraged YP to attend the session in advance, "Some staff did not knock on YP's doors until 11 am (when we were meeting!)" (F1).[5] Within the theme of practical challenges, facilitators discussed how changes in availability of regularly used delivery space presented issues in adhering to the delivery style when they had to deliver a session in the staff office. Furthermore, other practical barriers discussed included the scheduling of sessions, which were sometimes back-to-back sessions at different locations, and concerns regarding the time of the sessions. Finally, a variety of participant behaviours presented as a barrier, including poor punctuality, coming, and going throughout the session, fluctuations in attendance, inappropriate discussions, and use of mobile phones. For example, F2 discussed how participant behaviour provided a challenge to delivering the session, "the [number] of young people might not have actually been the direct challenge, but more the staggered arrival of them".

Two themes were consistent barriers across both phases of the programme. High support needs, including those related to physical and mental health, as well as language barriers and drug use, required increased one-to-one support for some YP. Meeting high support needs for the YP was emotionally draining for facilitators and meant that they could not always spend time with each YP during an activity. For example:

> "I think maybe the way I did strengths profiles could have been a bit better, like, maybe going around more and having some one-to-one conversations, I did try and do this a bit but was also helping YP2 with it as it was her first time doing it." (F2)

Inconsistencies in service provider staff taking a strengths-based or psychologically-informed (i.e., PIE) approach towards young people and engagement within sessions and activities was a further barrier common to both phases. F3 gave the example of a staff member encouraging young people to help tidy up more during Phase 2 by promoting external rewards, "She [staff member] then said there would be a prize for whoever did the best job. I didn't like this so much as it goes against the autonomy supportive approach we try to have." Offering an external reward goes against the delivery style encouraged in MST4Life™ which values development of intrinsic motivation where participants are encouraged to do things for the feeling it creates internally rather than an external reward (extrinsic motivation).

Table 5
Main themes highlighting barriers and enablers of fidelity of delivery style across phases 1 and 2 of MST4Life™.

Barriers and Enablers to fidelity of delivery style	Phase 1 (local service)	Phase 2 (OAE)
Barriers		
Poor communication	X	
High support needs of young people	X	X
Practical challenges	X	
Participant behaviour	X	
Departures from PIE/MST approach by service provider staff	X	X
Weather		X
Enablers		
Teamwork and communication	X	X
Positive participant behaviour	X	X
Outdoor instructor support		X

6.2.2. Enabling factors to fidelity of delivery style

Whilst no enabling factors were distinct to Phase 1, there were two themes which were present across both phases of the programme: teamwork and communication. These enablers occurred between facilitators, facilitators and service provider staff, and between service provider staff and YP. One facilitator described an occasion where a member of service provider staff was present in a Phase 1 session, as well as commenting on their behaviours to communicate with YP before the start of the session:

> "S1 was absolutely great not only in bringing the YP there, but also chasing up beforehand, ringing around other staff and also

[5] Note that throughout the qualitative results section letters and numbers have been used to protect the identity of facilitators, young people and service staff; F is used when the individual is the facilitator, YP when the individual is the young person, CD when the individual is a co-deliver, S when the individual is a member of staff from the homeless service, and O when the individual is an observer.

displaying co-deliverer behaviours in terms of helping the girls problem solve rather than giving them the answers." (F1)

Furthermore, another facilitator discussed in-depth an example of communication between facilitators and service provider staff. A young person who required high need support due to ill-health experienced a manifestation of their condition whilst out on the cave walk. The facilitator highlighted how excellent communication (e.g., clear speech, appropriate body language, and effective decision making) between all parties enabled the activity to continue running smoothly, safely, and enjoyably for the group:

"S2 was sat with her (YP), and (O1) was there too, as I walked over O1 and I made a subtle and efficient decision that I'd stay with YP and S2 and she (O1) went on with the rest of the group." (F2)

Co-delivery staff also discussed the importance of good communication and teamwork prior to the sessions themselves:

"I have been meeting F2 half hour before the session which really helps prepare for the session, F2 has generally been very helpful, encouraging, and his laid-back nature has helped me to start to flourish in my delivery without feeling too pressured." (CD1)

Positive participant behaviour was another enabling theme which was identified across both phases. Facilitators discussed YP showing positive behaviours including supporting one another generally, as well as during activities and engaging in activities even under difficult circumstances. One facilitator described how even in bad weather the YP "threw themselves" into the activity. Additionally, facilitators also discussed how even when the number of YP in the session was low (e.g., 2), that they "were very engaged so that wasn't a problem".

Distinct to Phase 2, the final theme of what enabled fidelity of delivery style was the support received from the instructors of the OAE centre and was frequently discussed by facilitators during their Phase 2 reflections. Communication between themselves (facilitators) and outdoor instructors, as well as the outdoor instructors' support of YP was described as consistent with the desired delivery style. For example, one facilitator discussed how the preparation and communication prior to the outdoor adventure trip was key to enabling the delivery of the project for one particular young person:

"One thing I feel like we did really well was prep for the needs of this trip (mainly YP1), which contributed towards what felt like a pretty smooth-running trip… The phone conversations with the outdoor instructors really helped too, so definitely keeping that communication going is key." (F1)

7. Discussion

This study examined the extent to which it is possible for frontline service staff of a housing service to deliver a psychologically informed intervention with high fidelity to delivery style through assessing and comparing fidelity of delivery style by deliverers and frontline service staff, and through exploring barriers and enablers to delivering MST4Life™ using the intended delivery style. Results show that MST4Life™ was delivered with high fidelity to delivery style. The reasons for this are presented in this discussion and considered within the context of the theories (PIEs and SDT) which underpin the programme's design as well as the wider fidelity literature.

7.1. Programme adherence

Results from this evaluation show that it is possible for frontline staff from a youth housing service to deliver psychologically informed programmes with high fidelity to delivery style. Although frontline staff did not score as highly as those with extensive psychology training (postgraduate level), they did deliver the programme with high fidelity, suggesting it is therefore possible to train frontline service staff to deliver psychologically informed interventions with high fidelity to delivery style. Frontline staff from the housing service also demonstrated a greater number of need thwarting behaviours and fewer need supportive behaviours than facilitators with psychological training. Research by Smith et al. (2016) found similar results, whereby despite scores which showed a high number of need supportive behaviours displayed by coaches, novice coaches displayed a higher number of need thwarting behaviours compared to those with more experience. Overall, the most experienced facilitators delivered MST4Life™ with higher levels of fidelity. Having more experienced facilitators delivering with higher levels of fidelity was expected from the outset. University facilitators initially delivered the programme; mid-way through the project, co-delivery training was subsequently introduced to enable frontline service staff to deliver the programme as part of the housing service's sustainability plans. Therefore, further improvements in fidelity to intended delivery style could be achieved by frontline staff as their experience of delivering the programme increases (Weiss & Westerhof, 2020; Wenz-Gross & Upshur, 2012).

Furthermore, reflective practice may also be a contributing factor in enabling the deliverers to deliver the sessions with such high fidelity. Reflective practice enables individuals (and teams) to become more self-aware, resilient, better able to cope with change, maintain and enhance skills and improve job performance (Edward & Hercelinskyj, 2007; Heyler, 2016). As such, despite delivering within a complex intervention and with participants who have complex needs, engagement in reflective practice encouraged deliverers to reflect on positive aspects of past delivery and areas for improvement and how that might be achieved. The deliverers from the university with postgraduate level psychology training and greater experience in programme delivery also had more experience engaging in regular self-initiated reflective practice compared with frontline staff. University deliverers kept a diary of delivery experiences from the programme outset and engaged in a debrief following each session as well as team reviews during and post programme delivery.

Frontline service staff did engage in self-reflection by completing the self-reflection forms after each session, which also provided a reminder of core delivery elements. Additionally, having an experienced facilitator model the desired delivery style during each session and engaging in informal post-session discussion debriefs with the university facilitator helped to clarify and extend knowledge of the co-deliverers beyond the initial training. This variation in prior experience of self-reflection may offer an explanation to the greater difference in self-rated scores compared to observed scores of the frontline staff who were also less experienced in delivery. The difference in delivery and self-reflection experience, as well as scores, emphasizes the importance of continuous staff training (on delivery and reflective practice) and reflective practice in enabling facilitators to improve and develop skills and styles required to deliver complex interventions in the community setting (McNicholas, Lennox, Woodcock, Bell, & Reed, 2019; Weiss & Westerhof, 2020).

Results show that despite differences in fidelity between the more experienced programme delivers and the co-deliverers, overall, MST4Life™ was delivered with high fidelity of delivery style. However, it is still important to highlight that as well as factors that enabled fidelity of delivery style, deliverers also experienced several barriers.

7.2. Factors affecting delivery style

Of the barriers and enablers to the fidelity of delivery style, a few were only relevant to certain phases. Most, however, related to both phases of the programme and highlight the challenges and key factors to delivering a complex intervention with high fidelity. Distinct to Phase 1 were practical challenges and participant behaviour. Practical challenges included difficulties with available space to deliver sessions, and the high volume of sessions within a short time period including

sometimes two sessions back-to-back but in different locations. Participant behaviour included lateness to sessions, poor attitude towards sessions, and mobile phone use. These behaviours are consistent with other complex interventions delivered to disadvantaged youth. Mihalic et al. (2008) cited common barriers as: (a) finding room in the schedule; (b) classroom management difficulties; and (c) gaining full support from key stakeholders. The latter was also perceived as a barrier to fidelity of delivery style by MST4Life™ facilitators across both phases and was also reported by Durlak and DuPre (2008). MST4Life™ facilitators discussed how even when they were trying to deliver in the correct style, sometimes the member of staff (both those directly co-delivering and those acting in a support role) from the service provider would not always display appropriate behaviours or communicate effectively (communication was inconsistent from frontline service staff and as such also appears within the enabling factors).

However, more commonly within this theme facilitators discussed feeling mixed levels of support and endorsement of the programme by some frontline service staff supporting the programme. For example, sometimes support staff seemed unhappy when a programme time (chosen by YP) runs past the time they are due to finish, leading to interrupting programme delivery to remind of the time, creating a sense of having to rush and feeling as though it was extra work encouraging YP to attend. University facilitators perceived that mixed support and endorsement from frontline staff led to young people arriving with no idea about the programme, as well as facilitators feeling pressured to rush sessions due to staff who were in a hurry to leave. Where support from frontline staff was mixed it limited available time for facilitators to engage the YP in in-depth discussions or reflections due to starting sessions late or having to explain additional things. This is an important barrier to consider. Firstly, it highlights the importance of factoring in time to engage in in-depth conversation, an opportunity during which facilitators can build rapport with individuals and create a sense of relatedness (Deci and Ryan, 2002), and begin to further understand and listen to a YP. Time to engage in in-depth discussions can allow the strengths-based approach to be implemented in a way that is tailored to the individual, enabling further discovery of a YPs strengths. These strengths can be used throughout the programme and beyond where a meaningful rationale is given for its application in the individual's wider life. Secondly, it highlights the broader challenges faced by frontline service staff including but not limited to, underfunded services, high caseloads, low-wages (and working second jobs) and family responsibilities (e.g., having to leave to pick up children). Staff play a crucial role in maintaining young people's engagement (Parry et al., 2021) in MST4Life™. It is vital to work collaboratively with frontline staff when making key decisions to ensure that these facilitate frontline staff in supporting the project and young people's engagement within it as well as being appropriate for the YP. This combination of findings further indicates the essential role frontline service staff play (as key gate keepers) in supporting effective programme delivery and is a vital element in bridging the research to applied practice gap (Ely, O, & Munroe-Chandler, 2020; Weiss & Westerhof, 2020).

As well as barriers to delivering MST4Life™ with fidelity to the intended delivery style, there were also several enabling factors. One was relevant only to Phase 2 of the programme, and that was the support and good communication of the outdoor instructors which enabled sessions to be prepared so that they would support the complex needs of the young people. Good communication with outdoor instructors was key during the outdoor activities, as it meant that sessions could be designed with an appropriate level of challenge, whilst remaining achievable for a variety of abilities: a key aspect for meeting young people's competency needs (Deci & Ryan, 2002). Similar to Mihalic and colleagues (2008), who also reported positive participant behaviour as enabling factor, MST4Life™ facilitators also perceived this to be the case, especially when participants were engaged and on time. For example, facilitators discussed how when good communication occurred between themselves and service provider staff, and between service provider staff and young people it ensured fidelity by helping with practical issues such as starting on time and consistency in need supportive behaviours.

In summary, MST4Life™ can be delivered with high fidelity, showing a greater level of fidelity than other studies as well as a lower discrepancy between observed and self-report scores, likely due to the extensive experience of the facilitators. There are a variety of perceived barriers and enablers to delivering with fidelity that are consistent with other complex interventions, including support from stakeholders, participant behaviour and practical considerations. The high fidelity of delivery style scores achieved in this study and barriers and enablers discussed tell us that it is possible to deliver programmes with high fidelity despite the complex settings in which they are delivered. Additionally, the barriers highlighted (logistical and practical challenges, l mixed levels of support from service staff and participant behaviour) are consistent with those from previous studies (Melde, Esbensen, and Tusinski (2006) and Mihalic et al. (2008). Despite progress in co-designing complex programmes since the early 2000's, the barriers faced are largely still the same and are linked to larger systemic changes required in funding housing services for disadvantaged youth.

7.3. Limitations and future research directions

Firstly, although training on using the observation tools and the context of MST4Life™ was provided to observers, and regular meetings were held to confirm that observers had developed the same interpretation for each behaviour, there was no opportunity to practice observations prior to data collection in the main study. Although not possible in the current study due to time constraints, the inclusion of video training would have strengthened observer training prior to data collection, through exposure to simulated sessions that depicted increasingly complex events as well as showing need supportive and need thwarting behaviours (Dempsey, Iwata, Fritz, & Rolider, 2012). The inclusion of videos showcasing a variety of situations as well as need supportive and need thwarting behaviours would better enable discussion around observers' understanding of the observation tool during training and throughout the data collection period, ensuring a greater level of accuracy and minimisation of observer drift during data collection (Yoder, Symons, & Lloyd, 2018). Inclusion of videos within training observers is especially important given the complex context of MST4Life™ where participant and logistical challenges mean that there can be a variety of influences which may make conducting observations more challenging.

A further limitation of this study and a vital step forward for future research is the need to evaluate staff training. This study evaluated delivery team performance (fidelity of delivery style); however, no rigorous evaluation was conducted of the training the frontline service staff received prior to delivery. Understanding effectiveness of staff training is key to illuminating present and future staff performance and ensuring sufficient and appropriate training is delivered to enable those delivering programmes to do so with high fidelity. Evaluating both fidelity of delivery as well as training received also enables more valuable provision of top-up training to address areas where staff would find more training beneficial. Evaluation of staff training was not included within this study due to time constraints, however it is something we strongly recommend future studies include as part of their process evaluations.

8. Lessons learned

Important lessons learned through this study can inform future program planning and evaluation in the area of complex community interventions:

8.1. Allowing for flexibility

Flexibility is not synonymous with lack of rigour in terms of programme delivery or evaluation. The programme was designed to enable flexibility in delivery by having the delivery style as a core active ingredient so that adaptations could be made to meet the needs of programme participants (Harn, Parisi, & Stoolmiller, 2013; Webster-Stratton, Reinke, Herman, & Newcomer, 2011). Ensuring adaptability to changing situations in an evaluation is vital to support data collection when evaluating complex community programmes; there are many factors that cannot be controlled, particularly when programme participants have high need support requirements (Cohen et al., 2008). On a small number of occasions programme participants did not turn up, leading to a session being cancelled, and no observation would take place due to the fixed nature of data collection time points in the pilot study. As a result, a range of data collection points were used in the main study. For example, data collection point 2 took place between sessions 4–6, and meant that data collection could take place in a systematic and rigorous way. Additionally, observing in an OAE setting requires a flexible approach to completion of data collection tools to account for safety (e.g., during climbing the "Old Man" at Coniston and navigating rocky terrain) and challenges of completing the form during poor weather conditions (e.g., rain). Requesting completion of the form by observers within 1 h of the activity meant that their own safety could be maintained when it was not safe or practical for the form to be completed in real time. It is vital that flexibility within programmes is reflected and continued within evaluations.

8.2. Meeting the emotional needs of the young people

Given the complex needs and vulnerable nature of the young people involved in the intervention, ensuring the young people felt comfortable was key to minimising the impact of the observer's presence. As such, during the design of the study, project developers and facilitators decided that the observer should build rapport with the young people rather than assume a more traditional observation style throughout (e.g., creating a sense of distance; Melde et al., 2006).

8.3. Facilitating frontline staff support expectations

Despite many advocates and high levels of support from frontline staff within the housing service, staff could sometimes display mixed levels of support towards the programme. As aforementioned, this mixed support was likely as result of not being involved in key decisions which affected their ability to support the delivery of MST4Life™ amongst high caseloads. It is vital to not only manage frontline staff expectations in terms of how to support programme delivery but to create an environment in which this also possible. How to create this environment and manage multiple needs (staff and participants) was a steep learning curve for both researchers and frontline service staff combined with having to meet the complex and varied needs of the YP accessing the service. As such, we recommend researchers collaborate with frontline staff and YP from the beginning (e.g., design) and throughout the programme when making key decisions to ensure that programme delivery (and support of) and YP participation is achievable in practice by all involved. When the programme is intended to be incorporated into the service long-term (as with MST4Life™) the importance of facilitating frontline staff support through collaboration is of even greater importance.

9. Conclusions

As the UK faces increasing numbers of youth experiencing homelessness, there is a pressing need for effective interventions that can both address the many and often co-occurring challenges that these young people experience and be delivered with high fidelity within the constraints faced by public services. This unique study was the first to assess fidelity of delivery style of a complex community intervention for youth experiencing homelessness, serving as an exemplar of the practices a program provider is meant to use in implementing a programme to ensure the context is strengths-based and psychologically informed. The study provides evidence that frontline staff can deliver psychologically informed programmes with high fidelity of delivery style and consequently, can be trained to deliver psychologically based programmes in complex settings. Furthermore, it highlights the vital nature of process evaluations in understanding mechanisms which make a programme successful (or not), such as flexible design and delivery and buy-in from service staff. This study goes some way to filling the need for greater understanding of how to implement programmes with high fidelity to delivery style as well as the feasibility to conducting such evaluations. However, it is essential that when researchers and organisations design programmes, they consider the evaluation of processes from the outset to ensure that programme outcomes can be correctly attributed to the programme, and to improve the uptake of successful programmes which will be of most benefit to their recipients.

Funding

This work was supported by funding from the Economic and Social Research Council, United Kingdom,through the lead authors studentship. The work was also supported by the registered charity St Basils, United Kingdom (registered charity number: 1080,154).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Anderson, L. M., Oliver, S. R., Michie, S., Rehfuess, E., Noyes, J., & Shemilt, I. (2013). Investigating complexity in systematic reviews of interventions by using a spectrum of methods. *Journal of Clinical Epidemiology, 66*(11), 1223–1229. https://doi.org/10.1016/j.jclinepi.2013.06.014

Balu, R., & Quint, J. (2014). Measuring implementation fidelity in success for all. In C. Meyers, & C. W. Brandt (Eds.), *Implemetation fidelity in education research: Designer and evaluator considerations* (Vol. 1, pp. 154–175). Routledge.

Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., , … Czajkowski, S., & Treatment Workgroup of the NIH Behaviour Change Consortium. (2004). Enhancing Treatment Fidelity in Health Behavior Change Studies: Best Practices and Recommendations From the NIH Behavior Change Consortium. *Health Psychology, 23*(5), 443–451. https://doi.org/10.1037/0278-6133.23.5.443

Bernard, R. (2018). Assessment of fidelity to the housing first principles of the HABITAT Programme. *European Journal of Homelessness, 12*(3).

Blundo, R. (2001). Learning strengths-based practice: Challenging our personal and professional frames. *Families in Society, 82*(3), 296–304.

Borrelli, B. (2011). The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *Journal of Public Health Dentistry, 71*(s1), S52–S63. https://doi.org/10.1111/j.1752-7325.2011.00233.x

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2013). *Successful qaulitative research: A guide for beginners*. SAGE.

Breitenstein, S. M., Gross, D., Garvey, C. A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation fidelity in community-based interventions. *Research in Nursing & Health, 33*(2), 164–174. https://doi.org/10.1002/nur.20373

Bruns, E. J., Burchard, J. D., Suter, J. C., Leverentz-Brady, K., & Force, M. M. (2004). Assessing fidelity to a community-based treatment for youth: The wraparound fidelity index. *Journal of Emotional and Behavioral Disorders, 12*(2), 79–89. https://doi.org/10.1177/10634266040120020201

Bumbarger, B. K., & Perkins, D. (2008). After randomised trials: Issues related to dissemination of evidence-based interventions. *Journal of Children's Services, 3*(2), 55–65.

Cohen, D. J., Crabtree, B. F., Etz, R. S., Balasubramanian, B. A., Donahue, K. E., Leviton, L. C., … Green, L. W. (2008). Fidelity versus flexibility: Translating evidence-based research into practice. *American Journal of Preventive Medicine, 35*(5, Supplement), S381–S389. https://doi.org/10.1016/j.amepre.2008.08.005

Collins, C., Lavender, M., Brown, S., Sheffield, P., & Andrew Aligne, C. (2013). Cyclopedia: sustaining a positive youth development program through community partnership. *International Journal on Disability and Human Development, 12*(4), 457–465. https://doi.org/10.1515/ijdhd-2013-0208

Cronley, C., & Evans, R. (2017). Studies of resilience among youth experiencing homelessness: A systematic review. *Journal of Human Behavior in the Social Environment, 27*(4), 291–311. https://doi.org/10.1080/10911359.2017.1282912

Cumming, J., Skeate, A., & Anderson, G. (2017). *St basils psychologically informed environments – Meeting the emotional and psychological needs of young homeless people*. L. a. I. N. Housing. ⟨https://stbasils.org.uk/files/2017-01-02/PIEHLIN_CaseStudy_130_StBasilsPIE_v01.pdf⟩.

Cumming, J., Whiting, R., Parry, B. J., Clarke, F. J., Holland, M. J., Cooley, S. J., & Quinton, M. L. (2022). The My Strengths Training for Life™ program: Rationale, logic model, and description of a strengths-based intervention for young people experiencing homelessness. *Evaluation and Program Planning, 91*. https://doi.org/10.1016/j.evalprogplan.2021.102045

Deci, E. L. (2000). The 'What' and 'Why' of goal pursuits: human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–269. https://doi.org/10.1207/S15327965PLI1104_01

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality, 19*(2), 109–134. https://doi.org/10.1016/0092-6566(85)90023-6

Deci, E. L., & Ryan, R. M. (2002). Self-determination research: Reflections and future directions. In Edward L Deci, & Richard M Ryan (Eds.), *Handbook of self-determination research* (pp. 431–441). University of Rochester Press.

Dempsey, C. M., Iwata, B. A., Fritz, J. N., & Rolider, N. U. (2012). Observer training revisited: A comparison of in vivo and video instruction. *Journal of Applied Behavior Analysis, 45*(4), 827–832. https://doi.org/10.1901/jaba.2012.45-827

Dobson, D., & Cook, T. (1980). Avoiding type III error in program evaluation: results from a field experiment. *Evaluation and Program Planning, 3*(4), 269–276. https://doi.org/10.1016/0149-7189(80)90042-7

DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. *American Journal of Community Psychology, 30*(2), 157–197. https://doi.org/10.1023/A:1014628810714

Edward, K.-L., & Hercelinskyj, G. (2007). Burnout in the caring nurse: Learning resilient behaviours. *The British Journal of Nursing, 16*(4), 240–242. https://doi.org/10.12968/bjon.2007.16.4.22987

Ely, F. O., O, J., & Munroe-Chandler, K. J. (2020). How intervention research designs may broaden the research-to-practice gap in sport psychology. *Journal of Sport Psychology in Action*, 1–13. https://doi.org/10.1080/21520704.2020.1798573

Greenhalgh, T., Jackson, C., Shaw, S., & Janamian, T. (2016). Achieving research impact through co-creation in community-based health services: Literature review and case study. *The Milbank Quarterly, 94*(2), 392–429. ⟨www.jstor.org/stable/24869175⟩.

Gwadz, M., Freeman, R., Leonard, N. R., Kutnick, A., Silverman, E., Ritchie, A., … Powlovich, J. (2019). Understanding organizations serving runaway and homeless youth: A multi-setting, multi-perspective qualitative exploration. *Child Adolesc Social Work J, 36*(2), 201–217. https://doi.org/10.1007/s10560-018-0563-3

Hardeman, W., Michie, S., Fanshawe, T., Prevost, A. T., McLoughlin, K., & Kinmonth, A. L. (2008). Fidelity of delivery of a physical activity intervention: Predictors and consequences. *Psychology & Health, 23*(1), 11–24. https://doi.org/10.1080/08870440701615948

Harn, B., Parisi, D., & Stoolmiller, M. (2013). Balancing fidelity with flexibility and fit: What do we really know about fidelity of implementation in schools. *Exceptional Children, 79*(3), 181–193. https://doi.org/10.1177/001440291307900204

Hasson, H. (2010). Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implementation Science: Islem-, 5*(1), 67. https://doi.org/10.1186/1748-5908-5-67

Hawe, P., Shiell, A., Riley, T., & Gold, L. (2004). Methods for exploring implementation variation and local context within a cluster randomised community intervention trial. *Journal of Epidemiology and Community Health, 58*(9), 788. https://doi.org/10.1136/jech.2003.014415

Heyler, R. (2016). *The work based learning student handbook* (2 ed.). Palgrave MacMillan.

Holt, N. L., Neely, K. C., Slater, L. G., Camiré, M., Côté, J., Fraser-Thomas, J., … Tamminen, K. A. (2017). A grounded theory of positive youth development through sport based on results from a qualitative meta-study. *International Review of Sport and Exercise Psychology, 10*(1), 1–49. https://doi.org/10.1080/1750984X.2016.1180704

Homeless Link. (2021). *Young and homeless*. H. Link. ⟨https://www.homeless.org.uk/sites/default/files/siteattachments/Young%20and%20Homeless%202021.pdf⟩.

Iachini, A. L., Beets, M. W., Ball, A., & Lohman, M. (2014). Process evaluation of "Girls on the Run": Exploring implementation in a physical activity-based positive youth development program. *Evaluation and Program Planning, 46*(1). https://doi.org/10.1016/j.evalprogplan.2014.05.001

Kenyon, D. B., McMahon, T. R., Simonson, A., Green-Maximo, C., Schwab, A., Huff, M., & Sieving, R. E. (2019). My journey: Development and practice-based evidence of a culturally attuned teen pregnancy prevention program for native youth. *International Journal of Environmental Research and Public Health, 16*(3), 470 (Article).

Kitzman-Ulrich, H., Saunders Ruth, P., Griffin, S., Wilson Dawn, K., Meyers Duncan, C., & Mansard, L. (2009). Using process evaluation for program improvement in dose, fidelity and reach: the ACT trial experience. *International Journal of Behavioral Nutrition and Physical Activity, 6*(1), 79. https://doi.org/10.1186/1479-5868-6-79

Krabbenborg, M., Boersma, S., & Wolf, J. R. (2013). A strengths based method for homeless youth: Effectiveness and fidelity of Houvast. *BMC Public Health, 13*(1), 359. https://doi.org/10.1186/1471-2458-13-359

Lorentson, M., Joo Oh, Y., & LaBanca, F. (2014). STEM21 digital academy fidelity of implementation: valuation and assessment of program components and implementation. In M. Coby V, & B. Christopher W (Eds.), *Implemetation fidelity in education research: designer and evaluator considerations* (Vol. 1, pp. 53–85). Routledge.

MacKie, P. K., & Thomas, I. (2014). Nations apart? Experiences of single homelessness across Great Britain. *CRISIS*. ⟨https://www.crisis.org.uk/ending-homelessness/homelessness-knowledge-hub/types-of-homelessness/nations-apart-experiences-of-single-homeless-people-across-great-britain-2014/⟩.

Maynard, B. R., Peters, K., Vaughn, M., & Sarteschi, C. (2013). Fidelity in after-school program intervention research: A systematic review. *Research on Social Work Practice, 23*(6), 613–624. https://doi.org/10.1177/1049731513491150

McNicholas, C., Lennox, L., Woodcock, T., Bell, D., & Reed, J. E. (2019). Evolving quality improvement support strategies to improve Plan–Do–Study–Act cycle fidelity: A retrospective mixed-methods study. *BMJ Quality & Safety, 28*(5), 356–365. https://doi.org/10.1136/bmjqs-2017-007605

McQuillin, S. D., & Lyons, M. D. (2021). A national study of mentoring program characteristics and premature match closure: The role of program training and ongoing support. *Prevention Science, 22*(3), 334–344. https://doi.org/10.1007/s11121-020-01200-9

McQuillin, S. D., Straight, G. G., & Saeki, E. (2015). Program support and value of training in mentors' satisfaction and anticipated continuation of school-based mentoring relationships. *Mentoring & Tutoring for Partnership in Learning, 23*(2), 133–148. https://doi.org/10.1080/13611267.2015.1047630

Melde, C., Esbensen, F.-A., & Tusinski, K. (2006). Addressing program fidelity using onsite observations and program provider descriptions of program delivery. *Evaluation Review, 30*(6), 714–740. https://doi.org/10.1177/0193841×06293412

Mihalic, S. F., Fagan, A. A., & Argamaso, S. (2008). Implementing the LifeSkills Training drug prevention program: Factors related to implementation fidelity. *Implementation Science, 3*(1). https://doi.org/10.1186/1748-5908-3-5

Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., … Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ: British Medical Journal, 350*, h1258. https://doi.org/10.1136/bmj.h1258

Parry, B. J., Quinton, M. L., Holland, M. J. G., Thompson, J. L., & Cumming, J. (2021). Improving outcomes in young people experiencing homelessness with My Strengths Training for Life™ (MST4Life™): A qualitative realist evaluation. *Children and Youth Services Review, 121*, Article 105793. https://doi.org/10.1016/j.childyouth.2020.105793

Pettee Gabriel, K. K., DiGioacchino DeBate, R., High, R. R., & Racine, E. F. (2011). Girls on the Run: a quasi-experimental evaluation of a developmentally focused youth sport program. *Journal of Physical Activity & Health, 8*, S285–S294.

Quinton, M. L., Clarke, F. J., Parry, B. J., & Cumming, J. (2021). An evaluation of My Strengths Training for LifeTM for improving resilience and well-being of young people experiencing homelessness. *Journal of Community Psychology, 49*(5), 1296–1314. https://doi.org/10.1002/jcop.22517

Rae, J., Samosh, J., Aubry, T., Tsemberis, S., Agha, A., & Shah, D. (2018). What helps and what hinders program fidelity to housing First: Pathways to housing DC. *European Journal of Homelessness, 12*(3).

Rixon, L., Baron, J., McGale, N., Lorencatto, F., Francis, J., & Davies, A. (2016). Methods used to address fidelity of receipt in health intervention research: a citation analysis and systematic review, 663-663 *BMC Health Services Research, 16*(1). https://doi.org/10.1186/s12913-016-1904-6.

Roth, J. L., & Brooks-Gunn, J. (2003). What exactly is a youth development program? Answers from research and practice. *Applied Developmental Science, 7*(2), 94–111. https://doi.org/10.1207/S1532480XADS0702_6

Roth, J. L., & Brooks-Gunn, J. (2016). Evaluating youth development programs: Progress and promise. *Applied Developmenta, 20*(3), 188–202. https://doi.org/10.1080/10888691.2015.1113879

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologis, 55*(1), 68–78. https://doi.org/10.1037//0003-066X.55.1.68

Sabalauskas, K. L., Ortolani, C. L., & McCall, M. J. (2014). Moving from pathology to possibility: Integrating strengths-based interventions in child welfare provision. *Child Care in Practice, 20*(1), 120–134. https://doi.org/10.1080/13575279.2013.847053

Schlosser, R. W. (2002). On the importance of being earnest about treatment integrity. *AAC: Augmentative & Alternative Communication, 18*(1), 36–45. ⟨https://www.tandfonline.com/doi/pdf/10.1080/aac.18.1.36.44?needAccess=true⟩.

Smith, B., & McGannon, K.R. (2018). Developing Rigor in Qualitative Research:Problems and Opportunities within Sport and Exercise Psychology.

Smith, N., Tessier, D., Tzioumakis, Y., Fabra, P., Quested, E., Appleton, P., … Duda, J. L. (2016). The relationship between observed and perceived assessments of the coach-created motivational environment and links to athlete motivation. *Psychology of Sport and Exercise, 23*, 51–63. https://doi.org/10.1016/j.psychsport.2015.11.001

Tidmarsh, G., Thompson, J. L., Quinton, M. L., & Cumming, J. (2022). Process evaluations of positive youth development programmes for disadvantaged young people: A systematic review [In press]. *Journal of Youth Development, 17*(2), 106–140. https://doi.org/10.5195/jyd.2022.1156

Vealey, S. R. (2007). Mental skills training in sport. In *Handbook of sport psychology / edited by Gershon Tenenbaum and Robert C. Eklund* (3rd ed. ed., Vol. 3, pp. 287–309). Hoboken, NJ: John Wiley, 2007.

Walton, H., Spector, A., Tombor, I., & Michie, S. (2017). Measures of fidelity of delivery of, and engagement with, complex, face-to-face health behaviour change interventions: A systematic review of measure quality. *British Journal of Health Psychology, 22*(4), 872–903. https://doi.org/10.1111/bjhp.12260

Webster-Stratton, C., Reinke, W. M., Herman, K. C., & Newcomer, L. L. (2011). The incredible years teacher classroom management training: The methods and principles that support fidelity of training delivery. *School Psychology Review, 40*(4), 509–529. https://doi.org/10.1080/02796015.2011.12087527

Weiss, L. A., & Westerhof, G. J. (2020). The happiness route: Finding alternatives to the problem-based approach in social work for vulnerable groups. *The Journal of Positive Psychology, 15*(5), 666–669. https://doi.org/10.1080/17439760.2020.1789713

Wenz-Gross, M., & Upshur, C. (2012). Implementing a primary prevention social skills intervention in urban preschools: Factors associated with quality and fidelity. *Early Education and Development, 23*(4), 427–450. https://doi.org/10.1080/10409289.2011.589043

Wierenga, D., Engbers, L. H., Van Empelen, P., Duijts, S., Hilderbrandt, V. H., & Van Mechelen, M. (2013). What is actually measured in process evaluations for worksite health promotion programs: A systematic review. *BMC Public Health, 13*(1), 1190–1191. https://doi.org/10.1186/1471-2458-13-1190

Yoder, P. J., Symons, F. J., & Lloyd, B. (2018). *Observational measurement of behaviour.* Brookes Publishing.

Grace Tidmarsh is a Doctoral Researcher in the School of Sport, Exercise and Rehabilitation Sciences at the University of Birmingham, UK. She is social scientist with interests relating to strengths-based sport psychology, process evaluation of interventions, and girls' engagement in physical activity/sport.

Richard Whiting is a Research Associate with a background in clinical psychological practice. He is a facilitator and evaluator for the MST4Life programme and associated projects in the community. Responsible for programme delivery, programme evaluation, developing implementation fidelity evaluation tools and processes, continuing improvement, staff training and building community partnerships to promote psychologically-informed working.

Janice L. Thompson, PhD, FACSM, FECSS, is Professor of Public Health Nutrition and Exercise at the University of Birmingham. She is a mixed-methods researcher with over 30 years of experience examining factors that influence diet- and physical activity-related lifestyle choices. She uses this information to develop and evaluate community-based culturally tailored lifestyle interventions. She is a Fellow of both the American College of Sports Medicine (ACSM) and the European College of Sport Science (ECSS). In 2015, she was awarded the ACSM Citation Award in recognition of her significant contributions to research and scholarship in exercise science.

Jennifer Cumming, PhD, CPsychol, AFBPsS, FHEA, is a Professor in Sport and Exercise Psychology at University of Birmingham. She is a community-engaged researcher who works collaboratively with third sector organisations and public services to co-develop and/or evaluate interventions to address important societal issues and translate knowledge into practice. Her programme of research investigates how individuals learn to effectively self-regulate and develop their mental skills for maximising potential and improving mental health, well-being, and social inclusion.