BMJ Open

Using mixed methods to assess treatment fidelity and its influencing factors in a complex self-management intervention

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-015452
Article Type:	Research
Date Submitted by the Author:	05-Dec-2016
Complete List of Authors:	Toomey, Elaine; National University of Ireland - Galway, School of Psychology Matthews, James; University College Dublin, Hurley, Deirdre; University College Dublin, Physiotherapy and Performance Science
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Sports and exercise medicine, Rheumatology, Qualitative research, Patient-centred medicine
Keywords:	implementation fidelity, PRIMARY CARE, complex interventions, behaviour change

Title

Using mixed methods to assess treatment fidelity and its influencing factors in a complex self-management intervention

Authors

Elaine Toomey[a]*, James Matthews[b], Deirdre A. Hurley[b].

[a]School of Psychology, Arts Millennium Building, National University of Ireland, Galway, Ireland

[b]School of Public Health, Physiotherapy and Sports Science, Health Sciences Centre, University College Dublin, Belfield, Dublin 4, Ireland

*Corresponding author

Corresponding author contact details

Email: elaine.toomey@nuigalway.ie

Phone: 00353-91494458

Word count

3853

ABSTRACT:

Objectives and design: Despite an increasing awareness of the importance of addressing treatment fidelity within complex behaviour change interventions, fidelity is often poorly assessed in such interventions with few examples of a comprehensive investigation of fidelity and its influencing factors. This mixed methods study aimed to establish the fidelity of a complex self-management intervention and explore the reasons for these findings using a concurrent/triangulation design.

Setting: Feasibility trial of the SOLAS self-management intervention for people with osteoarthritis and back pain, delivered in primary care physiotherapy

Methods and outcomes: 60 SOLAS intervention sessions were delivered across seven sites by nine physiotherapists. Fidelity of delivery was evaluated using 1) audio-recordings (n=60), direct-observations (n=24) and self-report checklists (n=60) and 2) individual semi-structured interviews with all physiotherapists (n=9). Quantitatively, fidelity was calculated using percentage means and standard deviations. Qualitative data were analysed using thematic analysis. Integration of quantitative and qualitative data occurred at an interpretation level using triangulation.

Results: Quantitatively, fidelity scores were high for all assessment methods; with self-report (92.7%) consistently higher than direct-observations (82.7%), or audio-recordings (81.7%). There was significant variation between physiotherapists' individual fidelity scores (69.8% - 100%). Both qualitative and quantitative data found that physiotherapists' knowledge and previous experience were factors that influenced their fidelity. The qualitative data also postulated participant-level and programme-level factors as additional elements that influenced fidelity.

Conclusion: This study contributes to the limited evidence regarding the use of fidelity assessment methods within the area of complex behaviour change interventions. The findings suggest a combination of quantitative methods is suitable for the assessment of fidelity within this context, depending on resources available. A mixed methods approach, integrating both quantitative and qualitative data, provided a more insightful understanding of fidelity and its

influencing factors, offsetting weaknesses inherent in using each research method in isolation.

Article summary – strengths and limitations

- This comprehensive investigation of fidelity and its influencing factors provides valuable information on fidelity assessment methods and factors to be considered in developing and evaluating complex behaviour change interventions
- The novel use of mixed methods to assess fidelity in this study enabled increased certainty in findings where qualitative data corroborated the quantitative results
- This study does not explore the fidelity of the *quality* of delivery (e.g. therapist competence) which will be reported elsewhere

USING MIXED METHODS TO ASSESS TREATMENT FIDELITY AND ITS INFLUENCING FACTORS IN A COMPLEX SELF-MANAGEMENT INTERVENTION

INTRODUCTION

Treatment fidelity is an important, yet often overlooked aspect of behaviour change interventions and has been defined as the degree to which an intervention, treatment or program is implemented as intended by the intervention developers. [1,2] Without adequately addressing fidelity in behaviour change research, it is uncertain that changes observed in study outcomes are due to the influence of the independent variable (the intervention being investigated) and not due to variability in its implementation, e.g. extraneous elements that may have been added (either accidentally or purposely), or essential elements of the intervention that were omitted. [3] In particular, as behaviour change interventions are often complex interventions that typically involve several components with the potential to affect or influence outcomes separately, it is especially important to incorporate adequate fidelity planning and assessment into the development of interventions of this nature. [4]

Despite a recent increased emphasis on improved assessment and reporting of what happens within complex behaviour change interventions, [5,6] fidelity is still poorly addressed within this context, with few examples of fidelity being addressed comprehensively or reported adequately. [7-9] Where studies have assessed fidelity within a behaviour change healthcare context, there is often limited exploration of the factors that might have influenced that fidelity. [10,11] Although the use of both quantitative and qualitative methods has been previously recommended to comprehensively assess fidelity, [12-14] this guidance is not consistently followed. Consequently, the use of quantitative methods in isolation may not allow for exploration of the factors influencing fidelity, knowledge of which could improve how fidelity is enhanced and assessed in future similar interventions. [15] For example, French et al. [11] used audio-recordings and self-report methods to assess the fidelity of delivery of an educational intervention for general practitioners. The authors acknowledged that the quantitative study design did not allow them to explore the reasons for variations in fidelity scores found.

According to existing guidelines for addressing fidelity within behaviour change research developed by the National Institutes of Health Behaviour Change Consortium (NIHBCC), [14,16] the fidelity of delivery of the intervention by providers (or Treatment Delivery) is one particularly important aspect of fidelity. Treatment delivery considers strategies that enhance the fidelity of delivery (e.g. using treatment manuals or intervention protocols) and methods that assess this delivery (e.g. provider self-report, audio or video-recorded observations and direct *in vivo* observations). However, although previous research has advocated a combination of these strategies in order to comprehensively assess fidelity, [17,18] limited examples exist within the literature. Additionally, few studies have explored the relationship between these methods, and the accuracy of potentially more feasible methods against the 'gold standard' of direct observations using pre-specified criteria [19] has been poorly investigated. [14,15,20] As a result, there is little evidence to justify the selection of one method over another, or to inform the use of multiple methods simultaneously.

The current study is set within the context of the SOLAS (Self-Management of Osteoarthritis and Low back pain through Activities and Skills) feasibility trial [ISRCTN49875385]. [21] The trial aims to test the feasibility of an intervention to promote self-management for people with osteoarthritis (OA) of the hip/knee and/or chronic low back pain (CLBP). The intervention consists of six weekly sessions of 90 minutes to be delivered by a primary care physiotherapist to a group of six to eight people. Each session is divided into education and exercise categories (each approximately 45 minutes in duration), with subcomponents as detailed in Table 1. Prior to participation, physiotherapists underwent a two-day training course where background variables were collected, in addition to a post-training assessment. [22] Details of the development and testing of the fidelity protocol used in this study are described elsewhere. [23]

The aim of this study was to explore and evaluate the fidelity of treatment delivery within the context of a complex behaviour change intervention using a mixed methods approach. Specifically, the study objectives were:

1) To evaluate the agreement of multiple methods for assessing fidelity of treatment delivery
2) To establish the fidelity of treatment delivery of the SOLAS complex behaviour change intervention
3) To explore the potential factors that may have influenced these fidelity results

Table 1: Structure, content and samples of SOLAS intervention						
Session Number	1	2	3	4	5	6
Self-management behaviours/ skills targeted	• Physical activity • Goal-setting	• Pacing (balancing activity/rest) • Physical activity • Goal-setting	• Balanced weight/ healthy diet • Physical activity • Goal-setting	• Managing pain with pain relief • Physical activity • Goal-setting	• Managing mood (with relaxation) • Physical activity • Goal-setting	• Long-term management • Physical activity • Goal-setting
Session Structure						
Category	**Category Subcomponents**	**Aim/content**				**Intended Duration**
Education	Materials	Participants are provided with materials intended to supplement and enhance their understanding and uptake of skills, such as pedometers, participant activity diaries, participant handbook and relaxation CDs				45 minutes
	Introduction and Review	At the start of each session the physiotherapist reviews goals and action plans with participants and problem-solving where necessary				
	Education	Physiotherapist facilitates a group discussion on the targeted self-management skill/behaviour of the session using Powerpoint slides and projector to provide information				

	Review and Planning	Before the session concludes, the physiotherapist briefly recaps participants' planned activity levels and action plans for the week ahead	
Exercise	Exercise	Participants are provided with an opportunity to attempt and practice a variety of exercises	45 minutes

Samples involved

	Wave 1				**Wave 2**						**Total n=**
Physiotherapy Site code	A	B	C	D	A	B	C	E	F	G	**10 (7)***
Physiotherapist code	A1	B1	C1	D1	A1	B2	C2	E1	F1	G1	**9****
Number of participants recruited	4	6	4	2	4	4	5	4	7	5	**45**
Sessions delivered	6	6	6	6	6	6	6	6	6	6	**60**
Direct Observations	3	3	3	3	2	2	3	2	1	2	**24 (40%)**
Self-Report	6	6	6	6	6	6	6	6	6	6	**60 (100%)**
Audio-Recordings (rater 1)	6	6	6	6	6	6	6	6	6	6	**60 (100%)**
Audio-Recordings (rater 2)	3	3	3	3	0	0	0	0	0	0	**12 (20%)**

**Seven sites were involved in total, but three of these delivered the intervention in both waves

*One physiotherapist delivered the intervention in both waves

METHODS

Ethics

Ethical approval for this study was granted by University College Dublin's Human Research Ethics Committee (LS-13-54). Written informed consent was obtained from all physiotherapists and participants prior to any observations or interviews.

Design

This observational study was a convergent/triangulation mixed methods design. [24] This mixed methods approach was chosen as it was felt that thorough integration of findings from both quantitative and qualitative methods would achieve a more comprehensive answer to the study questions by enabling the methods to be 'greater than just the sum of their parts'. [25] Figure 1 graphically illustrates the study design, outlining the sequence of research activities, the priority of the methods and the stage at which integration occurred. [26]

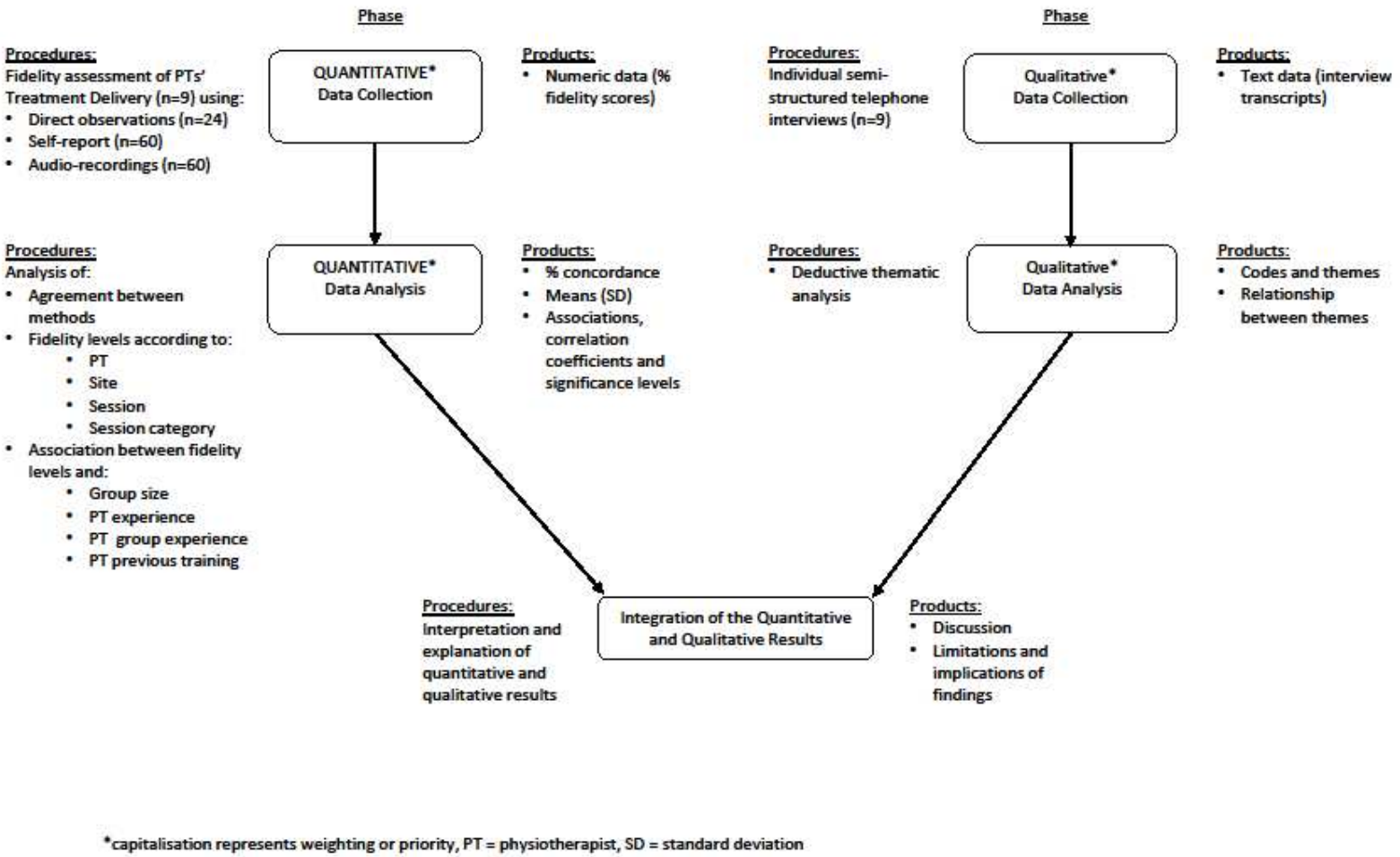

Figure 1: Diagrammatic representation of mixed methods convergent/triangulation design

Quantitative phase

Study sample and procedure:

Data were collected during two waves of the SOLAS feasibility trial (Table 1), representing 71% of the overall trial data. Following pilot testing, it was decided to conduct 24 (40% of sessions) randomly selected direct observations (rated by ET), 60 (100%) self-report (rated by the physiotherapists) and 60 (100%) audio-recordings (rated by ET) to assess fidelity of treatment delivery using a-priori checklists that had been previously found to be feasible for use. [23] To assess inter-rater reliability, 12 sessions (20%) were rated by a second independent rater (AK). [10] Checklists were structured into the SOLAS categories as detailed in Table 1 with components chosen based on the intended content. Session duration (dose) was documented by all methods, and attendance was recorded by self-report.

Data analysis:

Fidelity data analysis was consistent with standard procedures [15,27,28] using SPSS v20. Specifically, levels of agreement between methods and inter-rater reliability of audio-recorded data were assessed using percentage concordance. Overall mean fidelity levels and fidelity levels according to physiotherapy site, physiotherapist, session and session category were obtained by calculating total scores as a percentage of the total possible score. Means data were compared using ANOVAs and Kruskal-Wallis tests. Fidelity of dose was established by calculating the difference between the actual and the intended session duration using a one-sample Wilcoxon test. Levels of fidelity were interpreted as previously reported in the literature, with 80-100% adherence interpreted as 'high' fidelity, 51-79% as 'moderate' and 0-50% as 'low' fidelity. [3,29,30] Finally, the relationship between fidelity levels and 1) the number of participants present (group size) and 2) physiotherapist variables, i.e. experience (years qualified), group experience (years delivering group physiotherapy), knowledge of intervention (post-training evaluation score) and previous relevant training (Supplementary File 1), were calculated using Spearman's correlation coefficient and Mann-Whitney U test. These physiotherapist variables were chosen as previous research in other contexts has found that provider-level variables (e.g. experience) may influence fidelity of intervention delivery. [17,31-35]

Qualitative phase

The aim of the qualitative phase was to explore physiotherapists' opinions of fidelity of intervention delivery and the factors that they felt may have influenced their fidelity.

Individual semi-structured telephone interviews were conducted by an experienced qualitative researcher (SG) with each physiotherapist (n=9) within two weeks of intervention delivery completion. Interviews were audio-recorded and transcribed verbatim. Deductive thematic analysis was used to analyse the interviews as it is a flexible method that works with a range of research questions, including understanding people's experiences of programmes and health-care interventions. [36]

Meaningful units of text were highlighted within each interview, then summarised and coded. Codes dealing with similar issues were grouped across all interviews and refined into themes. The reliability of themes was established by a second reviewer (AK) who independently analysed a randomly selected sample of 50% of the transcript extracts using the coding framework. Percentage agreement was determined between the reviewers' respective coding of extracts. If agreement was less than 70%, consensus on conflicting decisions was obtained through discussion. [37]

Integration

Integration of quantitative and qualitative data occurred at an interpretation level using triangulation methodology. Specifically, a meta-matrix was created to facilitate comparison of the findings (Table 5). [38] This involved presenting the quantitative data in a tabular format alongside summarised qualitative themes, which enabled a transparent approach to determining convergence, discrepancy or silence across the findings of the data sets. [39] Convergence was defined as general agreement between the data sets on the element of comparison (e.g. overall quantitative fidelity score compared to the majority of physiotherapist opinions of their fidelity levels), while discrepancy was defined as general disagreement between the data sets on the element of comparison. [39] Silence was defined as where one set of results addressed a theme or example, but the other set of results did not yield any relevant data. [39]

RESULTS

Quantitative findings

Agreement:

Agreement between direct observations and self-report was 74.6%, 75.4% between self-report and audio-recordings (rater 1) and 86.6% between direct observations and audio-recordings (rater 1). Inter-rater reliability of audio-recordings (rater 1 vs rater 2) was 81.3%. Further detail is provided in Supplementary File 2.

Fidelity of content:

Fidelity was found to be high in all assessment methods, ranging from 81.7% (Audio-recordings) to 92.7% (Self-report). Table 2 details the fidelity results for each method with scores below 80 (cut-off for 'high' fidelity) shaded. Significant differences were found between Physiotherapist scores and between Category scores.

Table 2: Fidelity results calculated from % totals			
	Direct Observations (DO) % (SD)	**Self-Report (SR) % (SD)**	**Audio 1 (AO) % (SD)**
Total % mean fidelity score (SD)	82.7% (10)	92.7% (6.4)	81.7% (12)
% mean fidelity score per session (SD)* **			
1	88.8% (5.24)	95% (4.5)	91.6% (4.5)
2	82.8% (5.7)	92% (6.9)	86.8% (10.5)
3	85.6% (12.9)	96% (4.2)	81.4% (10.6)
4	83.3% (14.4)	90.9% (8.3)	75% (14)
5	74.1% (11.9)	89.4% (8.1)	78.7% (11)
6	82% (9.6)	92.7% (3.78)	74.9% (12)
% mean fidelity score per site (SD)* **			
A (delivered twice, same physiotherapist)	78.7% (7.6)	95% (5.4)	81.3% (11.5)
B (delivered twice, two physiotherapists)	76.7% (5.6)	92.8% (5.3)	71.1% (10)
C (delivered twice, two physiotherapists)	84.8% (11.8)	91.8% (7.7)	84.9% (8.1)
D	87.2% (4)	93.2% (2.9)	87.1% (4.4)
E	83.1% (13)	94.3% (3.8)	88.3% (8.7)

F	72.5%	85.2% (9.6)	72.9% (15)
G	100% (0)	94.7% (4.4)	92.8% (5.4)
% mean fidelity score per category (SD)*			
Materials	72.1% (19.4)	86% (17)	61.1% (29.6)
Introduction and Review	82.6% (16.3)	92.9% (12.8)	76.2% (24.5)
Education	93.3% (8.6)	97.1% (6.6)	95.4% (6.9)
Exercise	80.4% (14)	95.4% (7.1)	82.4% (13)
Review and Planning	77.1% (33)	90.8% (21.6)	69.8% (39.6)
% mean fidelity score per physio (SD)**			
A1	78.8% (7.6)	95.1% (5.4)	81.3% (11.5)
B1	76.1% (7.9)	92.2% (7.2)	72.3% (9.3)
B2	77.5% (0.4)	93.4% (3)	72.6% (12.5)
C1	93.4% (2)	85.2% (4.6)	91% (3.5)
C2	76.2% (11)	98.5% (1.9)	78.8% (6.6)
D1	87.2% (4)	93.2% (2.9)	87.1% (4.4)
E1	83.1% (13)	94.3% (3.8)	88.3% (8.7)
F1	72.5%	85.2% (9.6)	69.8% (14.7)
G1	100% (0)	94.7% (4.4)	95.8% (5.4)

*Significant differences between Categories according to DO (p=0.007), SR (p<0.001) and AO (p<0.001)

** Significant differences between Physiotherapists according to DO (p=0.019), SR (p=0.004) and AO (p<0.001)

***Significant differences between Sites (p<0.001) and between Sessions (p=0.007) according to AO, not significantly different

according to DO and SR.

Shaded areas = scores < 80%

Fidelity of dose (duration of sessions):

All methods found a statistically significant difference between the actual duration of the exercise component and its intended duration of 45 minutes. When the dosage was analysed for each individual session for all methods combined, a statistically significant difference ($p<0.001$) was only found for Session 1 between the actual and intended duration. The difference between the actual and intended dosage for all other sessions was not found to be statistically significant.

Group size – participants in attendance:

The average numbers of participants present for all groups across all sessions was 3, with an SD of 1.3. Overall, group attendance ranged from one to six participants for any session throughout both waves. The total number of participants recruited for each group was significantly different between physiotherapists (e.g. seven participants were recruited for the group delivered by physiotherapist F1 compared to only two recruited for the group delivered by D1), as were the numbers of participants present (average group size). The size of groups did not differ significantly between sessions.

Factors associated with fidelity:

Both direct observation and audio-recorded data showed a significant correlation between fidelity scores and the physiotherapists' post-training evaluation scores. Direct observation data also found a significant negative correlation between group sizes and the fidelity scores. Physiotherapist years qualified and experience of delivering groups were found to have significant, negative correlations with fidelity scores for the audio-recorded and self-report data respectively (Table 3).

Table 3: Association between fidelity scores and variables

Variable	Direct Observations	Self-Report	Audio-Recordings	Statistical Test
Group size (number present)	-0.434 (p=0.034)*	-0.215 (p=0.98)	-0.193 (p=0.151)	Spearman's *r* (p-value)
Physiotherapist experience (years qualified)	-0.09 (p=0.676)	-0.186 (p=0.154)	-0.346 (p=0.008)*	

Physiotherapist group experience (years delivering group physiotherapy)	0.171 (p=0.424)	-0.364 (p=0.004)**	0.136 (p=0.312)	
Physiotherapist Post-Training Evaluation Score (%)	0.581 (p=0.003)**	-0.152 (p=0.245)	0.314 (p=0.018)*	
Physiotherapist previous relevant training (yes/no)	*U*=33 (p=0.302)	*U*=201 (p=0.107)	*U* = 243 (p=0.840)	Mann-Whitney U (p-value)
*p is significant at p<0.05, **p<0.005				

Qualitative findings

Inter-rater reliability of coding achieved 81.6% agreement. Overall, physiotherapists felt that they had delivered the programme with good fidelity. All physiotherapists discussed some deviations from the protocol or adaptations made during delivery, i.e. goal-setting was found to be challenging to complete as intended. Other adaptations either concerned difficulties with use of programme materials (e.g. using the projector) as intended or providing additional information during the education content. Five physiotherapists also discussed deviation from protocol in relation to duration, mostly during the first session, with one stating that her *'time management around the education wasn't always exactly what it should have been'* (B1, transcript line 278-286). In terms of factors that influenced fidelity (i.e. reasons discussed for the aforementioned adaptations and deviations), six themes were found that were structured into three levels of factors – physiotherapist, participant and programme-level (Figure 2).

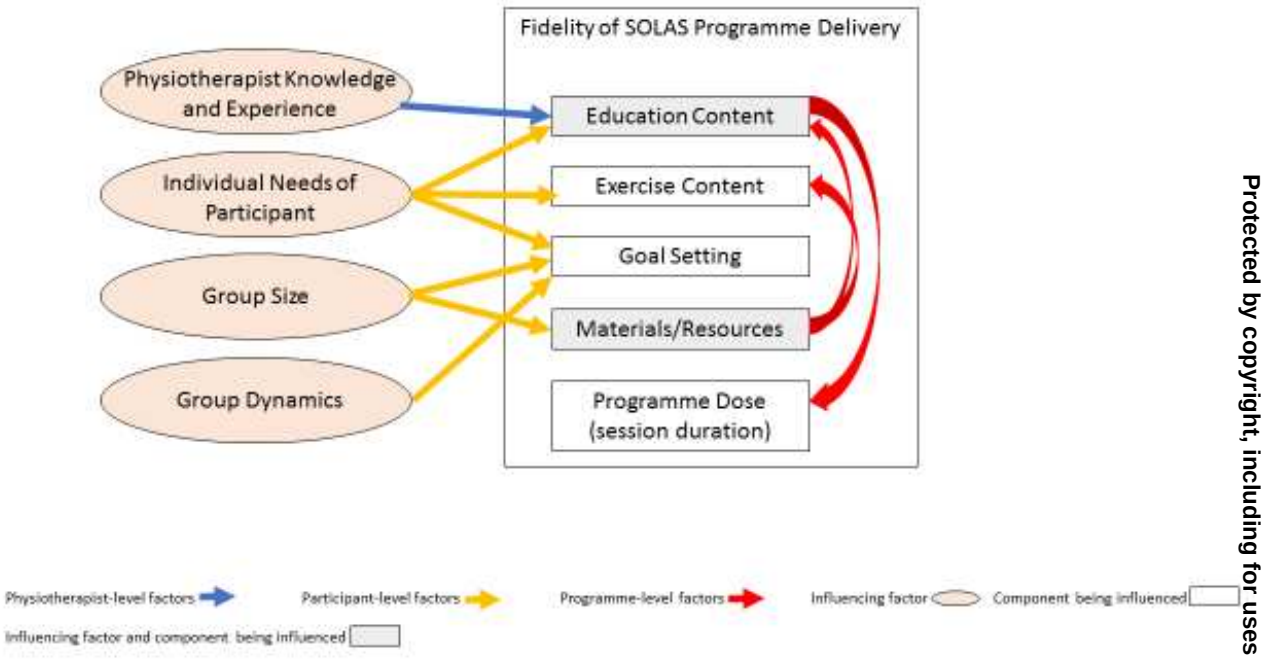

Figure 2: Visual representation of themes of qualitative interviews – factors influencing fidelity of SOLAS delivery

At the physiotherapist level, eight of the nine physiotherapists felt that knowledge of the content of the SOLAS programme facilitated their delivery of the education session and made it easier. Conversely, deviations from protocol within the education content discussed by six physiotherapists were due to the provision of additional information which was influenced by their previous experiences of delivering similar groups (e.g. talking more about pain pathophysiology because of previous classes delivered on this topic). This formed the theme '*Physiotherapist knowledge and experience influenced delivery of SOLAS - education content'.*

At the participant level, five physiotherapists felt that participants' individual needs such as their understanding of content or language literacy levels influenced the delivery of education and exercise components and that adaptation sometimes occurred in response to these needs, creating the theme of '*Individual needs influenced delivery of SOLAS - education, exercise, goal-setting'.* The number of participants present was discussed by seven physiotherapists as another participant-level factor that influenced fidelity of delivery and formed the participant-

level theme of '*Group size influenced delivery of SOLAS - goal-setting, use of materials'.* A further participant-level theme was *'Group dynamics influenced delivery of SOLAS - goal-setting'* as four physiotherapists felt that groups with good dynamics and interaction between participants led to better group discussions and better facilitation of goal setting.

The '*Amount of education content influenced delivery of SOLAS – dosage'* was a programme-level factor discussed by six physiotherapists, who felt that the amount of education content that was involved in the first session led to more time spent on the education aspect than intended as per protocol. Finally, all nine physiotherapists believed that the good resources (e.g. Powerpoint slides, venue) enhanced and facilitated the delivery of the programme as intended and that occasionally poor or problematic resources negatively influenced the delivery of the programme as intended. This created the theme '*Resources/materials influenced delivery of SOLAS – education and exercise content'.* Exemplary quotes are provided in Table 4.

Table 4: Qualitative findings of factors influencing fidelity results		
Factor level	**Theme**	**Exemplary quote (Physiotherapist code, transcript line)**
Physiotherapist	Physiotherapist knowledge and experience influenced delivery of SOLAS - education content	*'In my previous experience I would have done a lot more actually on the pain side of things…so in my previous class I would have had, you know, maybe one full class on maybe pain perception and, kind of, the influence of emotion and feelings…so I think I would have probably maybe talked a lot more around that pain section than maybe somebody else would have'.* (C1, 75-99)
Participant	Individual needs influenced delivery of SOLAS - education, exercise, goal-setting	'P*eople don't like writing them* [action plans] *there and then you know with pencils given and whatever - yes it's very hard to get people to write down things like that....Where I work there is a lot of people health literacy is very low…so therefore that's a challenge for them…so I tend to be very careful about pushing it out really'* (F1, 141-187).
Participant	Group size influenced delivery of SOLAS – goal-setting, use of materials	*'The only thing I might find a little bit hard would be the goal setting. I suppose you'd - that would be a bit more challenging because you'd have more numbers in the group*' (G1, 118-132)
Participant	Group dynamics influenced delivery of SOLAS - goal-setting	*'People were willing to engage you know as a group, in their goals.....so that made it very easy that we didn't actually have any clients that weren't willing to talk in the group, so it was very much an interactive group*' (E1, 225-231).
Programme	Amount of education content influenced delivery of SOLAS – dosage (session duration)	*'I found the content in week one was nearly too much... by the time I finished talking and ran through the exercises, the hour and a half was finished. And so nobody actually practiced any of the exercises on the first day'* (B2, 96-106).

Programme	Resources/materials influenced delivery of SOLAS – education and exercise content	*'The slides didn't work for me this time....You can't lock that room.... once or twice I didn't bring the laptop at all and I just had to print it out, all of the slides on A4 laminate and so we talked all the slides....'* (F1, 207-240) *'I think I only left out maybe three* [exercise] *stations, something like that. Because we didn't have a bouncer and...we didn't have a bed'* (C1, 113-121).

Integrating qualitative and quantitative findings

Each theme was further analysed according to the quantitative data comparing physiotherapists who scored 'high' (i.e. ≥80%; physiotherapists D1, C1, E1 and G1) to those who scored 'moderate' (i.e. ≥50-79%; B1, B2, C2 and F1). Physiotherapist A1 was not included in this analysis as her score was categorised as 'moderate' by direct observations and 'high' by audio-recorded data at 78.7% and 81.3% respectively. A difference between these physiotherapist groups was found in only one theme, '*Group size influenced delivery of SOLAS*'. Physiotherapists who scored higher on the fidelity assessments (average group size of 2.5 participants), believed it was easier to deliver goal-setting as intended with smaller groups. Conversely, physiotherapists with moderate fidelity scores (average group size of 3.7 participants) felt it was harder to facilitate goal-setting as intended with less numbers and believed it would be easier with bigger groups due to better engagement and group discussion. Further details of the triangulation are provided in Table 5.

Table 5: Triangulation of quantitative and qualitative results – convergence and discrepancy of findings

Outcome of interest	Quantitative fidelity findings		Qualitative Interview findings	Convergence/ Discrepancy/ Silence
Fidelity findings:	**Overall fidelity scores of SOLAS treatment delivery**	High fidelity (>80%) • Direct Observations: 82.7% • Self-Report: 92.7% • Audio-Recordings: 81.7%	Overall physiotherapists felt that their fidelity was good. Some adaptations and deviations were discussed as occurring in the delivery of the following aspects of the programme: • Goal-setting (Introduction and Review, Review and Planning categories) • Education content • Exercise content • Use of programme materials • Dosage of session 1 (session duration)	Convergence
	SOLAS categories scoring below 80% fidelity			
	Materials	Moderate fidelity (50-79%) • Direct Observations: 72.1% • Audio-Recordings: 61.1%		
	Introduction and Review	• Audio-recordings: 76.2%		
	Review and Planning	• Direct Observations: 77.1% • Audio-Recordings: 69.8%		
	Fidelity of dose – sessions significantly different to intended dose			
	Session 1	• Education duration: 58.9'* • Exercise duration: 31.4'*		
Factors influencing	**Correlation between quantitative variables and fidelity scores**		*Theme 1: Physiotherapist knowledge and experience influenced delivery of SOLAS - education content*	Convergence

fidelity findings:	Physiotherapist Experience (years qualified)	• Audio-recordings: -0.333 (p=0.011)***		
	Physiotherapist Group Experience (years delivering group physiotherapy)	• Self-report: -0.430 (p=0.018)***		
	Physiotherapist Post-Training Evaluation Score (%)	• Direct observations: 0.581 (p=0.003)** • Audio-recordings: 0.314 (p=0.018)***	*Theme 1: Physiotherapist knowledge and experience influenced delivery of SOLAS - education content*	Convergence
	Group size (average numbers of participants present)	• Direct observations: -0.434 (p=0.034)***	*Theme 3: Group size influenced delivery of SOLAS – goal-setting, use of materials*	Convergence
	No corresponding quantitative data		*Theme 2: Individual needs influenced delivery of SOLAS - education, exercise, goal-setting*	Silence
			Theme 4: Group dynamics influenced delivery of SOLAS - goal-setting	
			Theme 5: Amount of education content influenced delivery of SOLAS – dosage (session duration)	
			Theme 6: Resources/materials influenced delivery of	

		SOLAS – education and exercise content	
***p<0.001, **p<0.005, ***p<0.05**			

DISCUSSION

The aim of this mixed methods study was to explore and evaluate treatment fidelity within a feasibility trial of a complex behaviour change intervention using multiple assessment methods. The study found good agreement between researcher-delivered direct observation and audio-recorded fidelity assessment methods, with lower agreement found between provider self-report and researcher-delivered methods. The intervention content was delivered overall with high fidelity, with some variation between physiotherapists and between certain intervention categories. The intervention dose was found to have deviated significantly from intended during the first session only. Subsequently, qualitative interviews with physiotherapists confirmed these fidelity findings. Finally, both qualitative and quantitative data showed that physiotherapists' knowledge and previous experience, as well as the group size, were factors that influenced their fidelity of intervention delivery. The qualitative data contributed further, and postulated additional participant and programme-level factors as aspects that also influenced the overall fidelity results.

Agreement between direct observations and audio-recordings for assessing the fidelity of treatment delivery was found to be excellent. [40] Agreement between both of these methods and provider self-report assessment was lower, as providers consistently rated themselves higher than the independent raters. These findings are perhaps unsurprising, as both direct observations and audio-recordings were rated by the same researcher, and numerous previous studies have shown that providers' subjective assessments of fidelity are often rated higher than independent assessments. [27,28,41] Taking direct observations as the commonly-cited 'gold standard', [11,14,16] these findings reinforce that self-report methods may not be the most accurate method for assessing fidelity in a complex behaviour change study. However, they may still have their place for recording data and also for enhancing fidelity to the protocol by serving as an aide memoire for providers. [42] Although direct observations and audio-recordings have their own limitations [20,43], previous piloting of these assessment methods found that they were feasible and acceptable to physiotherapists. [23] Additionally, the good agreement between audio-recordings and direct observations found in this study suggests that audio-recordings may be a viable alternative with limited resources, as has been done in similar interventions. [44] However, where resources allow, a combination of multiple quantitative methods may provide the most comprehensive assessment of fidelity.

One of the key study findings of this study was reinforcing the value of using mixed methods research for the assessment of fidelity. This approach was emphasised in the recent MRC guidelines for conducting process evaluations of complex interventions [6], and is becoming increasingly used in the widespread implementation of evidence-based interventions [45] but does not yet appear to be common practice within fidelity assessments of behaviour change interventions. [10,17] The integration of quantitative and qualitative results enabled the triangulation of findings to provide a better overall picture of the fidelity of the SOLAS intervention and its influencing factors. The importance of the qualitative contribution to answering the 'why' question is evident in the fact that the physiotherapist interviews unearthed strong participant and programme-level factors associated with fidelity results that were not apparent from the quantitative data alone. Whilst this may be predominantly due to the focus of the quantitative analysis on physiotherapist-level variables which were chosen based on existing literature, the participant and programme-level factors identified by this analysis such as group dynamics or amount of programme content may have been difficult to quantitatively analyse to demonstrate association with fidelity results.

This study found that the factors that may influence the fidelity of an interventions' delivery can occur on three levels – provider, participant and programme. Where previous studies have explored factors that have influenced fidelity of intervention delivery, many have focused solely on provider-level factors, demonstrating associations between fidelity and factors such as provider training or skills. [15,17,31,35] The findings of this study have valuable implications for future studies that aim to assess and enhance fidelity of similar interventions as they indicate that planning for fidelity should include considering potential influencing factors at each of these three levels. These results are consistent with recent conclusions by Masterson-Algar et al. in a stroke rehabilitation setting [34] who found that investigating fidelity within clinical trials should also take the individual needs of patients into account, and also concur with the findings of an education-based intervention that found the most common reason for adaptation within intervention delivery was insufficient time. [46]

On the physiotherapist-level, better knowledge of the intervention content and structure was found to positively correlate with quantitative fidelity scores, with a causative link established via the qualitative investigation. This echoes previous findings by Huijg et al. who showed that physiotherapist skill level was one of the most important predictors of fidelity. [17] A more targeted approach to enhancing fidelity in future interventions may

therefore be warranted, such as identifying physiotherapists at higher risk of lower fidelity using post-training evaluation scores and employing more focused fidelity assessment of delivery or further training for these providers, [31] as has been previously employed in similar interventions. [10,44] The results of the study also showed that physiotherapists with more experience of certain aspects tended to emphasise these at the expense of delivering other components as comprehensively as intended in the protocol. These experience-based adaptations invoke the well-established issue of adaptation versus fidelity. For years, research has debated the concept of fidelity versus adaptation, with the case made for both strict fidelity and for modifying interventions. [47] A third view is that both fidelity and adaptation are essential, and achieving an appropriate balance between both can allow an intervention to maximise its effectiveness, while being generalizable and flexible enough to be implementable. [48,49] To achieve this, our fidelity checklists included components that encouraged elements of treatment individualisation (e.g. individualised feedback regarding exercises). However, it may be that these checklists still did not allow for enough individualisation within delivery, an aspect that should be considered by other researchers seeking to conduct similar fidelity assessments.

A limitation of this study was the timing of the interviews, which did not allow a 'pure' convergent/triangulation design. Typically, the qualitative and quantitative methods occur concurrently in this design, [24] however, they were scheduled to take place after physiotherapists had experienced delivery of an entire six-week SOLAS intervention. Although a sequential explanatory design [26] could have been used, interviews were conducted within two weeks of the intervention completion in order to minimise recall bias, and it was not feasible to complete the qualitative analysis beforehand. Finally, this study mostly focuses on the adherence of delivery (e.g. intervention content) and does not address the *quality* or competence of delivery of SOLAS, or how it was delivered (e.g. interpersonal or communication style of the physiotherapist), as this is being addressed in a separate publication.

CONCLUSIONS

In process evaluations and fidelity assessments of large scale complex interventions it is often recommended to complete and report the results of the fidelity assessment before the trial outcomes so as not to bias reporting. [50] Future work will investigate the relationship between this evaluation of fidelity and the SOLAS feasibility trial outcomes (analysis currently

underway), enabling a potentially more insightful and accurate interpretation of findings. This study also has valuable implications for further research and the overall science of treatment fidelity as it contributes much-needed information to the limited current evidence for the application of fidelity assessment methods within the area of complex behaviour change. The findings have demonstrated how multiple quantitative methods can be used to assess the fidelity of a complex behaviour change intervention, and that a combination of methods may be most suitable, depending on their acceptability and available resources. We have also shown how the use of a mixed methods approach, integrating both quantitative and qualitative data, provides a more insightful understanding of the factors influencing fidelity.

AUTHORS' CONTRIBUTION

ET conceptualized and designed the study and conducted the data collection and analysis. JM and DAH provided guidance on the design. DAH is the principal investigator of the SOLAS trial and obtained funding for this study within the trial. All authors provided methodological consultation and were involved in substantial contributions to reviewing of the drafts. All authors approved the final version of the manuscript.

ACKNOWLEDGEMENTS

This study was funded as part of Health Research Award HRA_HSR/2012/24 by the Health Research Board of Ireland. The funding source had no other involvement in this publication. The authors would like to thank the physiotherapists and participants for taking part in the study. We would also like to thank Alison Keogh for her help in the inter-rater reliability analysis of the audio-recorded data and physiotherapist interviews and Dr. Suzanne Guerin for facilitating the interviews.

CONFLICTS OF INTEREST

None

REFERENCES:

1. Gearing RE, El-Bassel N, Ghesquiere A, Baldwin S, Gillies J, Ngeow E. Major ingredients of fidelity: a review and scientific guide to improving quality of intervention research implementation. *Clin. Psychol. Rev.* 2011;31(1):79-88
2. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implementation Science* 2007;2(40):1-9
3. Borrelli B, Sepinwall D, Ernst D, et al. A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *J. Consult. Clin. Psychol.* 2005;73(5):852-60
4. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337(a1655):1-6
5. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348
6. Moore GF, Audrey S, Barker M, et al. *Process evaluation of complex interventions: Medical Research Council guidance*, 2015.
7. Hoffmann TC, Walker MF, Langhorne P, Eames S, Thomas E, Glasziou P. What's in a name? The challenge of describing interventions in systematic reviews: analysis of a random sample of reviews of non-pharmacological stroke interventions. *BMJ Open* 2015;5(11)
8. Schinckus L, Broucke S, Housiaux M. Assessment of implementation fidelity in diabetes self-management education programs: a systematic review. *Patient. Educ. Couns.* 2014;96(1):13-21
9. Toomey E, Currie-Murphy L, Matthews J, Hurley DA. Implementation fidelity of physiotherapist-delivered group education and exercise interventions to promote self-management in people with osteoarthritis and chronic low back pain: A rapid review Part II. *Man. Ther.* 2015;20(2):287-94
10. Mars T, Ellard D, Carnes D, Homer K, Underwood M, Taylor SJC. Fidelity in complex behaviour change interventions: a standardised approach to evaluate intervention integrity. *BMJ Open* 2013;3(11)
11. French SD, Green SE, Francis JJ, et al. Evaluation of the fidelity of an interactive face-to-face educational intervention to improve general practitioner management of back pain. *BMJ Open* 2015;5(7)
12. Mowbray CT, Holter MC, Teague GB, Bybee D. Fidelity Criteria: Development, Measurement, and Validation. *American Journal of Evaluation* 2003;24(3):315-40
13. Poltawski L, Norris M, Dean S. Intervention fidelity: Developing an experience-based model for rehabilitation research. *J. Rehabil. Med.* 2014;46(7):609-15
14. Bellg AJ, Borrelli B, Resnick B, et al. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol.* 2004;23(5):443-51
15. Lorencatto F, West R, Christopherson C, Michie S. Assessing fidelity of delivery of smoking cessation behavioural support in practice. *Implementation Science* 2013;8(40)
16. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J. Public Health Dent.* 2011;71(s1):S52-S63
17. Huijg JM, Dusseldorp E, Gebhardt WA, et al. Factors associated with physical therapists' implementation of physical activity interventions in The Netherlands. *Phys. Ther.* 2015;95(4):539-57
18. McKenna JW, Flower A, Ciullo S. Measuring Fidelity to Improve Intervention Effectiveness. *Intervention in School and Clinic* 2014;50(1):15-21

19. Resnicow K, Davis M, Smith M, et al. How best to measure implementation of school health curricula: a comparison of three measures. *Health Education Research* 1998;13(2):239-50
20. Breitenstein SM, Gross D, Garvey CA, Hill C, Fogg L, Resnick B. Implementation fidelity in community-based interventions. *Res. Nurs. Health* 2010;33(2):164-73
21. Hurley DA, Hall AM, Currie-Murphy L, et al. Theory-driven group-based complex intervention to support self-management of osteoarthritis and low back pain in primary care physiotherapy: protocol for a cluster randomised controlled feasibility trial (SOLAS). *BMJ Open* 2016;6(1):e010728
22. Hurley DA, Murphy LC, Hayes D, et al. Using intervention mapping to develop a theory-driven, group-based complex intervention to support self-management of osteoarthritis and low back pain (SOLAS). *Implementation Science* 2016;11(1):1-29
23. Toomey E, Matthews J, Guerin S, Hurley DA. Development of a Feasible Implementation Fidelity Protocol Within a Complex Physiotherapy-Led Self-Management Intervention. *Phys. Ther.* 2016
24. Creswell JW, Plano Clark VL. *Designing and Conducting Mixed Methods Research*. 2 ed. Thousand Oaks, CA: SAGE Publications, 2011.
25. Barbour RS. The case for combining qualitative and quantitative approaches in health services research. *J. Health Serv. Res. Policy* 1999;4(1):39-43
26. Ivankova NV, Creswell JW, Stick SL. Using Mixed-Methods Sequential Explanatory Design: From Theory to Practice. *Field Methods* 2006;18(1):3-20
27. Hardeman W, Michie S, Fanshawe T, Prevost AT, McLoughlin K, Kinmonth AL. Fidelity of delivery of a physical activity intervention: Predictors and consequences. *Psychol. Health* 2008;23(1):11-24
28. Breitenstein SM, Fogg L, Garvey C, Hill C, Resnick B, Gross D. Measuring Implementation Fidelity in a Community-Based Parenting Intervention. *Nursing research* 2010;59(3):158-65
29. Perepletchikova F, Kazdin AE. Treatment Integrity and Therapeutic Change: Issues and Research Recommendations. *Clinical Psychology: Science and Practice* 2005;12(4):365-83
30. Garbacz L, Brown D, Spee G, Polo A, Budd K. Establishing Treatment Fidelity in Evidence-Based Parent Training Programs for Externalizing Disorders in Children and Adolescents. *Clin Child Fam Psychol Rev* 2014;17(3):230-47
31. Wang B, Stanton B, Deveaux L, et al. Factors influencing implementation dose and fidelity thereof and related student outcomes of an evidence-based national HIV prevention program. *Implement Sci* 2015;10(1):44
32. Langford R, Bonell C, Jones H, Campbell R. Obesity prevention and the Health promoting Schools framework: essential components and barriers to success. *Int. J. Behav. Nutr. Phys. Act.* 2015;12:15
33. Perepletchikova F, Hilt LM, Chereji E, Kazdin AE. Barriers to implementing treatment integrity procedures: survey of treatment outcome researchers. *J. Consult. Clin. Psychol.* 2009;77(2):212-8
34. Masterson-Algar P, Burton CR, Rycroft-Malone J, Sackley CM, Walker MF. Towards a programme theory for fidelity in the evaluation of complex interventions. *Journal of Evaluation in Clinical Practice* 2014;20(4):445-52
35. Taylor CA, Shaw RL, Dale J, French DP. Enhancing delivery of health behaviour change interventions in primary care: a meta-synthesis of views and experiences of primary care nurses. *Patient Educ. Couns.* 2011;85(2):315-22
36. Braun V, Clarke V. Using thematic analysis in psychology. *Qual. Res. Psychol.* 2006;3(2):77-101
37. Guerin S, Hennessy E. Pupils' definitions of bullying. *Eur J Psychol Educ* 2002;17(3):249-61
38. Wendler MC. Triangulation using a meta-matrix. *J. Adv. Nurs.* 2001;35(4):521-5
39. Farmer T, Robinson K, Elliott SJ, Eyles J. Developing and implementing a triangulation protocol for qualitative health research. *Qualitative Health Research* 2006;16(3):377-94

40. Fleiss J. Reliability of Measurement. In: Fleiss J, ed. The design and analysis of clinical experiments. New York: Wiley, 1986:1-32.
41. Carroll K, Nich C, Rounsaville B. Utility of Therapist Session Checklists to Monitor Delivery of Coping Skills Treatment for Cocaine Abusers. *Psychotherapy Research* 1998;8(3):307-20
42. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Publ Health Dentistry* 2011;71:S52 - S63
43. Leventhal H, Friedman MA. Does establishing fidelity of treatment help in understanding treatment efficacy? Comment on Bellg et al. (2004). *Health Psychol.* 2004;23(5):452-6
44. Pincus T, Anwar S, McCracken LM, et al. Delivering an Optimised Behavioural Intervention (OBI) to people with low back pain with high psychological risk; results and lessons learnt from a feasibility randomised controlled trial of Contextual Cognitive Behavioural Therapy (CCBT) vs. Physiotherapy. *BMC Musculoskelet. Disord.* 2015;16:147
45. Palinkas LA, Aarons GA, Horwitz S, Chamberlain P, Hurlburt M, Landsverk J. Mixed Method Designs in Implementation Research. *Administration and Policy in Mental Health* 2011;38(1):44-53
46. Hill LG, Maucione K, Hood BK. A focused approach to assessing program fidelity. *Prev Sci* 2007;8(1):25-34
47. Dane AV, Schneider BH. Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin. Psychol. Rev.* 1998;18(1):23-45
48. Castro FG, Barrera M, Jr., Martinez CR, Jr. The cultural adaptation of prevention interventions: resolving tensions between fidelity and fit. *Prev Sci* 2004;5(1):41-5
49. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology* 2008;41(3-4):327-50
50. Oakley A, Strange V, Bonell C, Allen E, Stephenson J. Health services research: process evaluation in randomised controlled trials of complex interventions. *BMJ* 2006;332:413

Supplementary Files

Supplementary File 1: Physiotherapist baseline characteristics										
	Wave 1				Wave 2					
Physiotherapy Site	A	B	C	D	A	B	C	E	F	G
Physiotherapist	A1	B1	C1	D1	A1	B2	C2	E1	F1	G1
Gender	F	F	F	M	F	F	F	F	F	F
Experience (years qualified)	5	25	8	12	5	25	10	6	19	11
Group experience (years delivering group physiotherapy)	1	3	8	2	1	10	3	1	15	7
Post-training evaluation score (%)	74%	81.5%	90.5%	83.4%	74%	88.5%	81.8%	81.2%	78.1%	94.4%
Previous training in similar interventions (Y/N)	N	Y	Y	Y	N	Y	Y	Y	Y	Y
F= female, M=male, Y=yes, N=no										

Supplementary File 2: Agreement between assessment methods				
	Direct Observation v Self-Report (n=24)	Direct Observation v Audio-Recordings (n=24)	Audio-Recordings v Self-Report (n=60)	Audio-Recordings Rater 1 v Rater 2 (Inter-rater reliability) (n=12)
Overall agreement	74.6%	86.6%	75.4%	84.6%
% Agreement per category:				
Materials	74.5%	82.67%	70.1%	84.6%
Introduction and Review	65.3%	86.5%	57%	81.6%
Education	84.6%	90.3%	87.3%	76.7%
Exercise	70.8%	86%	78%	83.3%
Review and Planning	50%	76.2%	46%	100%

BMJ Open

Using mixed methods to assess fidelity of delivery and its influencing factors in a complex self-management intervention

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-015452.R1
Article Type:	Research
Date Submitted by the Author:	10-Feb-2017
Complete List of Authors:	Toomey, Elaine; National University of Ireland - Galway, School of Psychology Matthews, James; University College Dublin, Hurley, Deirdre; University College Dublin, Physiotherapy and Performance Science
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Sports and exercise medicine, Rheumatology, Qualitative research, Patient-centred medicine
Keywords:	implementation fidelity, PRIMARY CARE, complex interventions, behaviour change

Title

Using mixed methods to assess fidelity of delivery and its influencing factors in a complex self-management intervention

Authors

Elaine Toomey[a]*, James Matthews[b], Deirdre A. Hurley[b].

[a]School of Psychology, Arts Millennium Building, National University of Ireland, Galway, Ireland

[b]School of Public Health, Physiotherapy and Sports Science, Health Sciences Centre, University College Dublin, Belfield, Dublin 4, Ireland

*Corresponding author

Corresponding author contact details

Email: elaine.toomey@nuigalway.ie

Phone: 00353-91494458

Word count

4077

ABSTRACT:

Objectives and design: Despite an increasing awareness of the importance of fidelity of delivery within complex behaviour change interventions, it is often poorly assessed. This mixed methods study aimed to establish the fidelity of delivery of a complex self-management intervention and explore the reasons for these findings using a convergent/triangulation design.

Setting: Feasibility trial of the Self-management of Osteoarthritis and Low back pain through Activity and Skills (SOLAS) intervention, delivered in primary care physiotherapy.

Methods and outcomes: 60 SOLAS sessions were delivered across seven sites by nine physiotherapists. Fidelity of delivery of pre-specified intervention components was evaluated using 1) audio-recordings (n=60), direct-observations (n=24) and self-report checklists (n=60) and 2) individual interviews with physiotherapists (n=9). Quantitatively, fidelity scores were calculated using percentage means and standard deviations of components delivered. Associations between fidelity scores and physiotherapist variables were analysed using Spearman's correlations. Interviews were analysed using thematic analysis to explore potential reasons for fidelity scores. Integration of quantitative and qualitative data occurred at an interpretation level using triangulation.

Results: Quantitatively, fidelity scores were high for all assessment methods; with self-report (92.7%) consistently higher than direct-observations (82.7%), or audio-recordings (81.7%). There was significant variation between physiotherapists' individual scores (69.8% - 100%). Both qualitative and quantitative data (from physiotherapist variables) found that physiotherapists' knowledge (Spearman's association at p=0.003) and previous experience (p=0.008) were factors that influenced their fidelity. The qualitative data also postulated participant-level (e.g. individual needs) and programme-level factors (e.g. resources) as additional elements that influenced fidelity.

Conclusion: The intervention was delivered with high fidelity. This study contributes to the limited evidence regarding fidelity assessment methods within complex behaviour change interventions. The findings suggest a combination of quantitative methods is suitable for the

assessment of fidelity of delivery. A mixed methods approach provided a more insightful understanding of fidelity and its influencing factors.

Article summary – strengths and limitations

- This mixed methods investigation of fidelity of delivery and its influencing factors provides valuable information on fidelity assessment methods and factors to be considered in developing and evaluating complex behaviour change interventions
- The novel use of mixed methods to assess fidelity in this study enabled increased certainty in findings where qualitative data corroborated the quantitative results
- This study does not explore the fidelity of the *quality* of delivery (e.g. therapist competence) or specific behaviour change techniques (BCTs) which will be reported in a separate publication

USING MIXED METHODS TO ASSESS FIDELITY OF DELIVERY AND ITS INFLUENCING FACTORS IN A COMPLEX SELF-MANAGEMENT INTERVENTION

INTRODUCTION

Fidelity of delivery is an important, yet often overlooked aspect of behaviour change interventions and has been defined as the degree to which an intervention, treatment or program is delivered as intended by the intervention developers. [1,2] Without adequately addressing fidelity in behaviour change research, it is uncertain that changes observed in study outcomes are due to the influence of the independent variable (the intervention being investigated) and not due to variability in its implementation, e.g. extraneous elements that may have been added (either accidentally or purposely), or essential elements of the intervention that were omitted. [3] In particular, as behaviour change interventions are often complex interventions that typically involve several components with the potential to affect or influence outcomes separately, it is especially important to incorporate adequate fidelity planning and assessment into the development of interventions of this nature. [4]

Despite a recent increased emphasis on improved assessment and reporting of what happens within complex behaviour change interventions, [5,6] fidelity is still poorly addressed within this context, with few examples of fidelity being assessed comprehensively or reported adequately. [7-9] Where studies have assessed fidelity within a behaviour change healthcare context, there is often limited exploration of the factors that might have influenced that fidelity. [10,11] Previous work that *has* specifically examined influencing factors in areas of public health, obesity and stroke research found that provider-level variables, such as experience, knowledge or skills, may influence fidelity of delivery. [12-15] Although the use of both quantitative and qualitative methods has been previously recommended to comprehensively assess fidelity, [16-18] this guidance is not consistently followed. Consequently, the use of quantitative methods in isolation may not allow for exploration of the factors influencing fidelity, knowledge of which could improve how fidelity is enhanced and assessed in future similar interventions. [19] For example, French et al. [11] used audio-recordings and self-report methods to assess the fidelity of delivery of an educational

intervention for general practitioners. The authors acknowledged that the quantitative study design did not allow them to explore the reasons for variations in fidelity scores found.

According to existing guidelines for addressing fidelity within behaviour change research developed by the National Institutes of Health Behaviour Change Consortium (NIHBCC), [18,20] the fidelity of delivery of the intervention by providers is one particularly important aspect of fidelity. This aspect considers strategies that enhance the fidelity of delivery (e.g. using treatment manuals or intervention protocols) and methods that assess this delivery (e.g. provider self-report, audio or video-recorded observations and direct *in vivo* observations). However, although previous research has advocated a combination of these strategies in order to assess fidelity in-depth, [21,22] limited examples exist within the literature. Additionally, few studies have explored the relationship between these methods, and the accuracy of potentially more feasible methods against the 'gold standard' of direct observations using pre-specified criteria [23] has been poorly investigated. [18,19,24] As a result, there is little evidence to justify the selection of one method over another, or to inform the use of multiple methods simultaneously.

The current study is set within the context of the SOLAS (Self-Management of Osteoarthritis and Low back pain through Activities and Skills) feasibility trial [ISRCTN49875385]. [25] The trial aims to evaluate the feasibility of providing the SOLAS intervention (experimental group) to promote self-management for patients with osteoarthritis (OA) of the hip/ knee and/or chronic low back pain (CLBP) compared to usual physiotherapy, which will serve as the pragmatic control group in order to determine the feasibility of moving to a full scale trial by following the MRC guidelines. The intervention consists of six weekly sessions of 90 minutes to be delivered by a primary care physiotherapist to a group of six to eight people. Each session is divided into education and exercise sections (each approximately 45 minutes in duration), and further split into categories as detailed in an intervention manual [25,26] (summarised in Table 1). Prior to participation, physiotherapists were provided with the manual during a two-day training course where background variables were collected, in addition to a post-training assessment. [26] Details of the development and testing of the fidelity protocol used in this study are described elsewhere. [27] This study also does not explore the fidelity of the *quality* of delivery (e.g. therapist competence) or specific behaviour change techniques (BCTs) which will be reported elsewhere.

The aim of this study was to establish the fidelity of delivery of a complex behaviour change intervention and the reasons for these findings using a mixed methods approach. Specifically, the study objectives were:

1) To evaluate the agreement of multiple methods for assessing fidelity of delivery
2) To establish the fidelity of delivery of the SOLAS complex behaviour change intervention
3) To explore the potential factors that may have influenced these fidelity results

Table 1: Structure, content and samples of SOLAS intervention						
Session Number	1	2	3	4	5	6
Self-management behaviours/ skills targeted	• Physical activity • Goal-setting	• Pacing (balancing activity/rest) • Physical activity • Goal-setting	• Balanced weight/ healthy diet • Physical activity • Goal-setting	• Managing pain with pain relief • Physical activity • Goal-setting	• Managing mood (with relaxation) • Physical activity • Goal-setting	• Long-term management • Physical activity • Goal-setting
Session Structure						
Section	**Category**	**Aim/content**				**Intended Duration**
Education	Materials	Participants are provided with materials intended to supplement and enhance their understanding and uptake of skills, such as pedometers, participant activity diaries, participant handbook and relaxation CDs				45 minutes
	Introduction and Review	At the start of each session the physiotherapist reviews goals and action plans with participants and problem-solving where necessary				
	Education	Physiotherapist facilitates a group discussion on the targeted self-management skill/behaviour of the session using Powerpoint slides and projector to provide information				

	Review and Planning	Before the session concludes, the physiotherapist briefly recaps participants' planned activity levels and action plans for the week ahead	
Exercise	Exercise	Participants are provided with an opportunity to attempt and practice a variety of exercises	45 minutes

Samples involved

	Wave 1				**Wave 2**						**Total n=**
Physiotherapy Site code	A	B	C	D	A	B	C	E	F	G	**10 (7)***
Physiotherapist code	A1	B1	C1	D1	A1	B2	C2	E1	F1	G1	**9****
Number of participants recruited	4	6	4	2	4	4	5	4	7	5	**45**
Sessions delivered	6	6	6	6	6	6	6	6	6	6	**60**
Direct Observations	3	3	3	3	2	2	3	2	1	2	**24 (40%)**
Self-Report	6	6	6	6	6	6	6	6	6	6	**60 (100%)**
Audio-Recordings (rater 1)	6	6	6	6	6	6	6	6	6	6	**60 (100%)**
Audio-Recordings (rater 2)	3	3	3	3	0	0	0	0	0	0	**12 (20%)**

**Seven sites were involved in total, but three of these delivered the intervention in both waves

*One physiotherapist delivered the intervention in both waves

METHODS

Ethics

Ethical approval for this study was granted by University College Dublin's Human Research Ethics Committee (LS-13-54). Written informed consent was obtained from all physiotherapists and participants prior to any observations or interviews.

Design

This observational study was a convergent/triangulation mixed methods design. [28] This mixed methods approach was chosen as it was felt that thorough integration of findings from both quantitative and qualitative methods would achieve a more comprehensive answer to the study questions by enabling the methods to be 'greater than just the sum of their parts'. [29] Figure 1 graphically illustrates the study design, outlining the sequence of research activities, the priority of the methods and the stage at which integration occurred. [30]

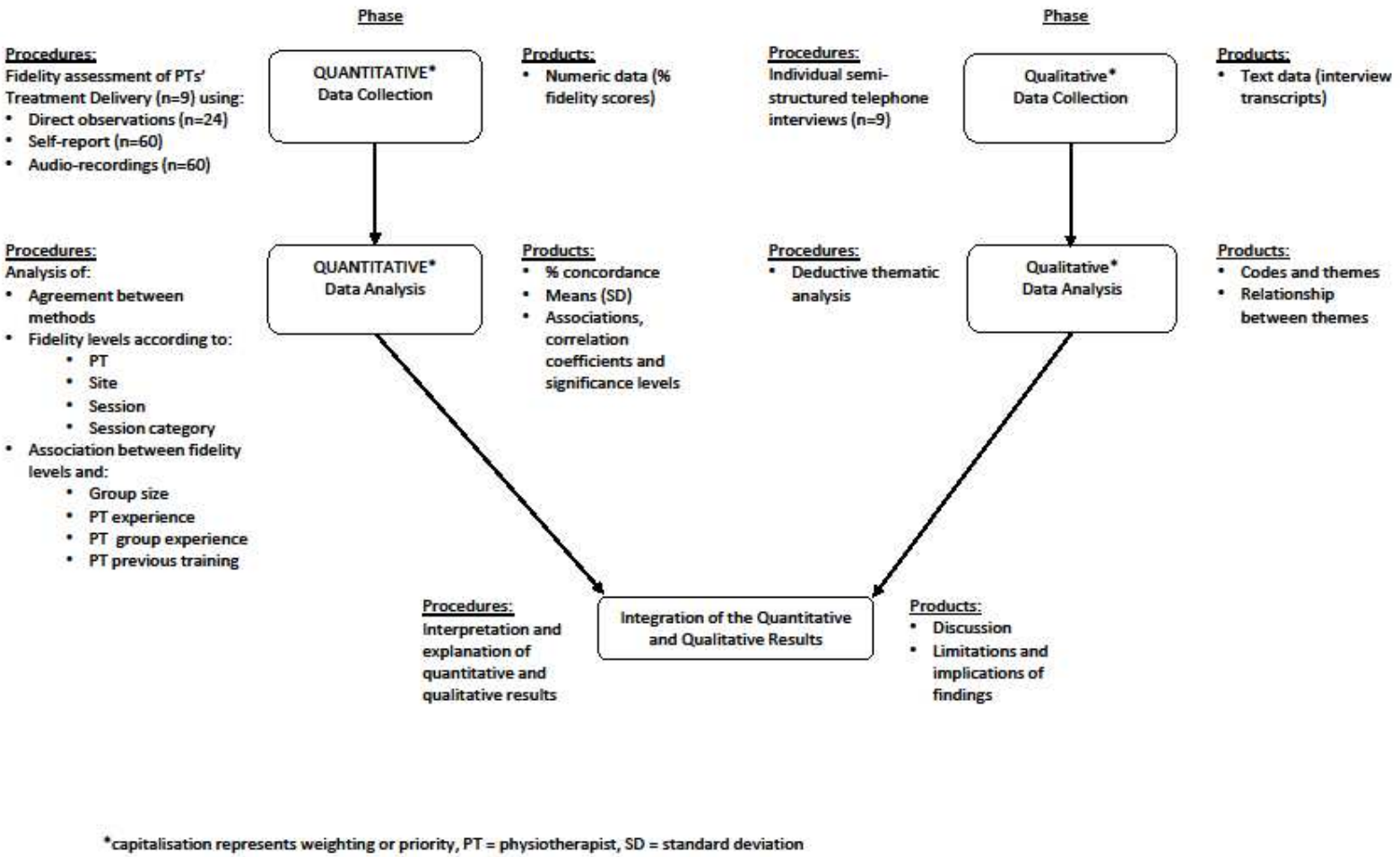

Figure 1: Diagrammatic representation of mixed methods convergent/triangulation design

Quantitative phase

Study sample and procedure:

Data were collected during two of the three waves of the SOLAS feasibility trial (Table 1), representing 71% of the overall trial data. Fidelity of delivery in this study refers to the assessment of both the delivery of session content, i.e. providers deliver the session categories and components as intended (summarised in Table 1) (fidelity of content) and session duration, i.e. providers deliver the session as long as intended (fidelity of duration). Following pilot testing, it was decided to conduct 24 (40% of sessions) randomly selected direct observations (rated by ET), 60 (100%) self-report (rated by the physiotherapists) and 60 (100%) audio-recordings (rated by ET) to assess fidelity of delivery using a-priori checklists (Supplementary File 1) that had been previously found to be feasible for use. [27] To assess inter-rater reliability, 12 sessions (20%) were rated by a second independent rater (AK). [10] Checklists consisted of approximately 25 components for each session, structured according to the SOLAS categories as detailed in Table 1. Components for each session were chosen to address each element specified in the SOLAS intervention manual (summarised in Table 1) [25,26] to be delivered during that session. Each component was rated as 'Yes/Present' equating to a score of two points, 'No/Absent' (zero points), or 'Attempted' (one point). Session duration was documented by all methods, and attendance was recorded by self-report.

Data analysis:

Fidelity data analysis was consistent with standard procedures [19,31,32] using SPSS v20. Specifically, levels of agreement between methods and inter-rater reliability of audio-recorded data were assessed using percentage concordance. Overall mean fidelity of content scores (i.e. percentage of manual-specified components delivered as intended) and fidelity scores according to physiotherapy site, physiotherapist, session and session category were obtained by calculating total actual scores (delivered components) as a percentage of the total possible score (intended components). Means data were compared using ANOVAs and Kruskal-Wallis tests. Fidelity of duration was established by calculating the difference between the actual and the intended session duration using a one-sample Wilcoxon test. Levels of fidelity were interpreted as previously reported in the literature, with 80-100% adherence interpreted as 'high' fidelity, 51-79% as 'moderate' and 0-50% as 'low' fidelity. [3,33,34] Finally, the relationship between fidelity scores and 1) the number of participants present (group size) and 2) physiotherapist variables, i.e. experience (years qualified), group

experience (years delivering group physiotherapy), knowledge of intervention (post-training evaluation score) and previous relevant training (Supplementary File 2), were calculated using Spearman's correlation coefficient and Mann-Whitney U test. These physiotherapist variables were chosen for reasons described in the introduction.

Qualitative phase

The aim of the qualitative phase was to explore physiotherapists' opinions of fidelity of intervention delivery and the factors that they felt may have influenced their fidelity. Individual semi-structured telephone interviews were conducted by an experienced qualitative researcher (SG) with each physiotherapist (n=9) within two weeks of intervention delivery completion. A topic guide with specific questions and probes related to fidelity was developed by the corresponding author (ET) (Supplementary File 3). Interviews were audio-recorded and transcribed verbatim. Deductive thematic analysis was used to analyse the interviews as it is a flexible method that works with a range of research questions, including understanding people's experiences of programmes and health-care interventions. [36]

Meaningful units of text were highlighted within each interview, then summarised and coded. Codes dealing with similar issues were grouped across all interviews and refined into themes. The reliability of themes was established by a second reviewer (AK) who independently analysed a randomly selected sample of 50% of the transcript extracts using the coding framework. Percentage agreement was determined between the reviewers' respective coding of extracts. If agreement was less than 70%, consensus on conflicting decisions was obtained through discussion. [37]

Integration

Integration of quantitative and qualitative data occurred at an interpretation level using triangulation methodology. Specifically, a meta-matrix was created to facilitate comparison of the findings (Table 5). [38] This involved presenting the quantitative data in a tabular format alongside summarised qualitative themes, which enabled a transparent approach to determining convergence, discrepancy or silence across the findings of the data sets. [39] Convergence was defined as general agreement between the data sets on the element of comparison (e.g. overall quantitative fidelity score compared to the majority of physiotherapist opinions of their fidelity levels), while discrepancy was defined as general

disagreement between the data sets on the element of comparison. [39] Silence was defined as where one set of results addressed a theme or example, but the other set of results did not yield any relevant data. [39]

RESULTS

Quantitative findings

Agreement:

Agreement between direct observations and self-report was 74.6%, 75.4% between self-report and audio-recordings (rater 1) and 86.6% between direct observations and audio-recordings (rater 1). Inter-rater reliability of audio-recordings (rater 1 vs rater 2) was 81.3%. Further detail is provided in Supplementary File 4.

Fidelity of content:

Fidelity was found to be high in all assessment methods, with a mean score of 81.7% (range of 61.1% - 95.8%) for the audio-recordings, 92.7% (85.2% - 96%) for the self-report and 82.7% (72.1% - 100%) for the direct observations. Table 2 details the fidelity results for each method with scores below 80 (cut-off for 'high' fidelity) shaded. Significant differences between physiotherapists' individual fidelity scores were found. Fidelity scores were also found to differ significantly according to the session category (e.g. the category 'Materials' was delivered with significantly less fidelity than the 'Education' category).

Table 2: Fidelity results calculated from % totals			
	Direct Observations (DO) % (SD)	**Self-Report (SR) % (SD)**	**Audio 1 (AO) % (SD)**
Total % mean fidelity score (SD)	82.7% (10)	92.7% (6.4)	81.7% (12)
% mean fidelity score per session (SD)* **			
1	88.8% (5.24)	95% (4.5)	91.6% (4.5)
2	82.8% (5.7)	92% (6.9)	86.8% (10.5)
3	85.6% (12.9)	96% (4.2)	81.4% (10.6)
4	83.3% (14.4)	90.9% (8.3)	75% (14)
5	74.1% (11.9)	89.4% (8.1)	78.7% (11)
6	82% (9.6)	92.7% (3.78)	74.9% (12)
% mean fidelity score per site (SD)* **			
A (delivered twice, same physiotherapist)	78.7% (7.6)	95% (5.4)	81.3% (11.5)
B (delivered twice, two physiotherapists)	76.7% (5.6)	92.8% (5.3)	71.1% (10)
C (delivered twice, two physiotherapists)	84.8% (11.8)	91.8% (7.7)	84.9% (8.1)
D	87.2% (4)	93.2% (2.9)	87.1% (4.4)
E	83.1% (13)	94.3% (3.8)	88.3% (8.7)

F	72.5%	85.2% (9.6)	72.9% (15)
G	100% (0)	94.7% (4.4)	92.8% (5.4)
% mean fidelity score per category (SD)*			
Materials	72.1% (19.4)	86% (17)	61.1% (29.6)
Introduction and Review	82.6% (16.3)	92.9% (12.8)	76.2% (24.5)
Education	93.3% (8.6)	97.1% (6.6)	95.4% (6.9)
Exercise	80.4% (14)	95.4% (7.1)	82.4% (13)
Review and Planning	77.1% (33)	90.8% (21.6)	69.8% (39.6)
% mean fidelity score per physio (SD)**			
A1	78.8% (7.6)	95.1% (5.4)	81.3% (11.5)
B1	76.1% (7.9)	92.2% (7.2)	72.3% (9.3)
B2	77.5% (0.4)	93.4% (3)	72.6% (12.5)
C1	93.4% (2)	85.2% (4.6)	91% (3.5)
C2	76.2% (11)	98.5% (1.9)	78.8% (6.6)
D1	87.2% (4)	93.2% (2.9)	87.1% (4.4)
E1	83.1% (13)	94.3% (3.8)	88.3% (8.7)
F1	72.5%	85.2% (9.6)	69.8% (14.7)
G1	100% (0)	94.7% (4.4)	95.8% (5.4)

*Significant differences between Categories according to DO (p=0.007), SR (p<0.001) and AO (p<0.001)

** Significant differences between Physiotherapists according to DO (p=0.019), SR (p=0.004) and AO (p<0.001)

***Significant differences between Sites (p<0.001) and between Sessions (p=0.007) according to AO, not significantly different

according to DO and SR.

Shaded areas = scores < 80%

Fidelity of duration:

All methods found a statistically significant difference between the actual duration of the exercise component and its intended duration of 45 minutes (Supplementary File 5). When this was analysed for each individual session for all methods combined, a statistically significant difference ($p<0.001$) was only found for Session 1 between the actual and intended duration. The difference between the actual and intended duration for all other sessions was not found to be statistically significant ($p>0.05$).

Group size – participants in attendance:

The average numbers of participants present for all groups across all sessions was 3, with an SD of 1.3 (Supplementary File 6). Overall, group attendance ranged from one to six participants for any session across both waves. The total number of participants recruited for each group was significantly different between physiotherapists (e.g. seven participants were recruited for the group delivered by physiotherapist F1 compared to only two recruited for the group delivered by D1), as were the numbers of participants present (average group size). The size of groups did not differ significantly between sessions.

Factors associated with fidelity:

Both direct observation and audio-recorded data showed a significant correlation between fidelity scores and the physiotherapists' post-training evaluation scores. Direct observation data also found a significant negative correlation between group sizes and the fidelity scores. Physiotherapist years qualified and experience of delivering groups were found to have significant, negative correlations with fidelity scores for the audio-recorded and self-report data respectively (Table 3).

Table 3: Association between fidelity scores and variables

Variable	Direct Observations	Self-Report	Audio-Recordings	Statistical Test
Group size (number present)	-0.434 (p=0.034)*	-0.215 (p=0.98)	-0.193 (p=0.151)	Spearman's *r* (p-value)
Physiotherapist experience (years qualified)	-0.09 (p=0.676)	-0.186 (p=0.154)	-0.346 (p=0.008)*	

Physiotherapist group experience (years delivering group physiotherapy)	0.171 (p=0.424)	-0.364 (p=0.004)**	0.136 (p=0.312)	
Physiotherapist Post-Training Evaluation Score (%)	0.581 (p=0.003)**	-0.152 (p=0.245)	0.314 (p=0.018)*	
Physiotherapist previous relevant training (yes/no)	*U*=33 (p=0.302)	*U*=201 (p=0.107)	*U* = 243 (p=0.840)	Mann-Whitney U (p-value)
*p is significant at p<0.05, **p<0.005				

Qualitative findings

Inter-rater reliability of coding achieved 81.6% agreement. Overall, physiotherapists felt that they had delivered the programme with good fidelity. All physiotherapists discussed some deviations from the protocol or adaptations made during delivery, i.e. goal-setting was found to be challenging to complete as intended. Other adaptations either concerned difficulties with use of programme materials (e.g. using the projector) as intended or providing additional information during the education content. Five physiotherapists also discussed deviation from protocol in relation to duration, mostly during the first session, with one stating that her *'time management around the education wasn't always exactly what it should have been'* (B1, transcript line 278-286). In terms of factors that influenced fidelity (i.e. reasons discussed for the aforementioned adaptations and deviations), six themes were found that were structured into three levels of factors – physiotherapist, participant and programme-level (Figure 2).

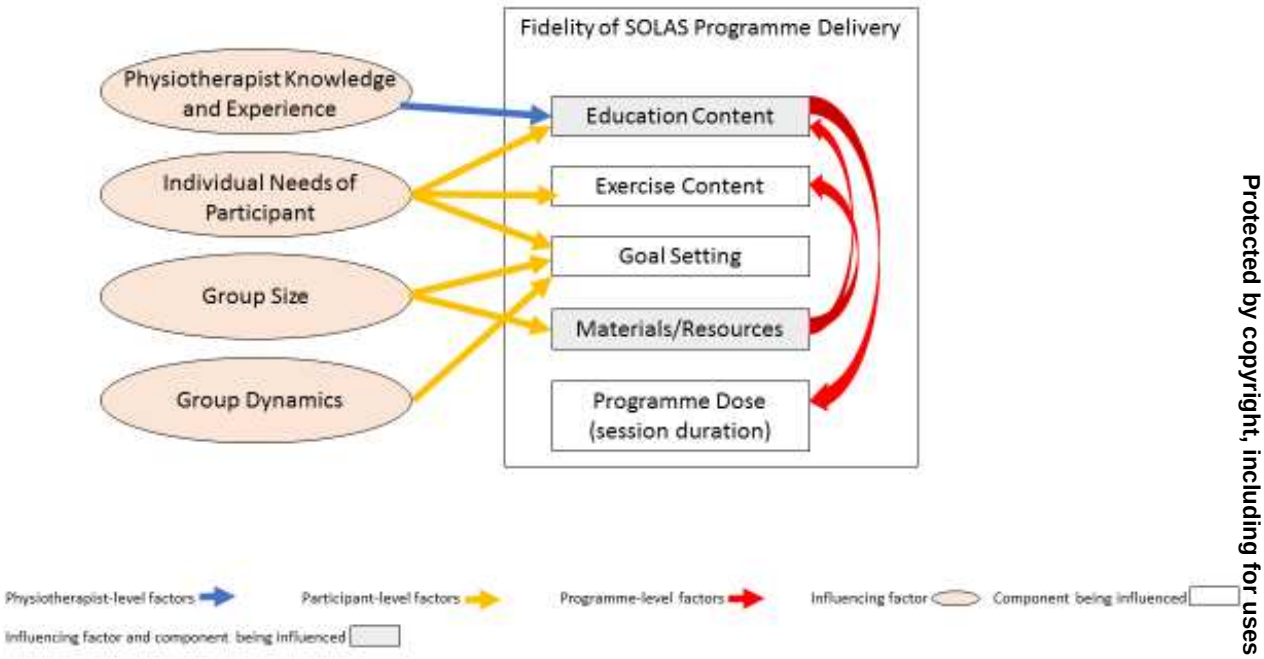

Figure 2: Visual representation of themes of qualitative interviews – factors influencing fidelity of SOLAS delivery

At the physiotherapist level, eight of the nine physiotherapists felt that knowledge of the content of the SOLAS programme facilitated their delivery of the education session and made it easier. Conversely, deviations from protocol within the education content discussed by six physiotherapists were due to the provision of additional information which was influenced by their previous experiences of delivering similar groups (e.g. talking more about pain pathophysiology because of previous classes delivered on this topic). This formed the theme '*Physiotherapist knowledge and experience influenced delivery of SOLAS - education content'.*

At the participant level, five physiotherapists felt that participants' individual needs such as their understanding of content or language literacy levels influenced the delivery of education and exercise components and that adaptation sometimes occurred in response to these needs, creating the theme of '*Individual needs influenced delivery of SOLAS - education, exercise, goal-setting'.* The number of participants present was discussed by seven physiotherapists as another participant-level factor that influenced fidelity of delivery and formed the participant-

level theme of '*Group size influenced delivery of SOLAS - goal-setting, use of materials'.* A further participant-level theme was *'Group dynamics influenced delivery of SOLAS - goal-setting'* as four physiotherapists felt that groups with good dynamics and interaction between participants led to better group discussions and better facilitation of goal setting.

The '*Amount of education content influenced delivery of SOLAS – duration* was a programme-level factor discussed by six physiotherapists, who felt that the amount of education content that was involved in the first session led to more time spent on the education aspect than intended as per protocol. Finally, all nine physiotherapists believed that the good resources (e.g. booklets and handouts, venue space) enhanced and facilitated the delivery of the programme as intended and that occasionally poor or problematic resources (e.g. lack of venue security) negatively influenced the delivery of the programme as intended. This created the theme '*Resources/materials influenced delivery of SOLAS – education and exercise content'.* Exemplary quotes are provided in Table 4.

Table 4: Qualitative findings of factors influencing fidelity results		
Factor level	**Theme**	**Exemplary quote (Physiotherapist code, transcript line)**
Physiotherapist	Physiotherapist knowledge and experience influenced delivery of SOLAS - education content	*'In my previous experience I would have done a lot more actually on the pain side of things…so in my previous class I would have had, you know, maybe one full class on maybe pain perception and, kind of, the influence of emotion and feelings…so I think I would have probably maybe talked a lot more around that pain section than maybe somebody else would have'.* (C1, 75-99)
Participant	Individual needs influenced delivery of SOLAS - education, exercise, goal-setting	'P*eople don't like writing them* [action plans] *there and then you know with pencils given and whatever - yes it's very hard to get people to write down things like that....Where I work there is a lot of people health literacy is very low…so therefore that's a challenge for them…so I tend to be very careful about pushing it out really'* (F1, 141-187)
Participant	Group size influenced delivery of SOLAS – goal-setting, use of materials	*'The only thing I might find a little bit hard would be the goal setting. I suppose you'd - that would be a bit more challenging because you'd have more numbers in the group*' (G1, 118-132)
Participant	Group dynamics influenced delivery of SOLAS - goal-setting	*'People were willing to engage you know as a group, in their goals.....so that made it very easy that we didn't actually have any clients that weren't willing to talk in the group, so it was very much an interactive group*' (E1, 225-231)
Programme	Amount of education content influenced delivery of SOLAS – duration	*'I found the content in week one was nearly too much... by the time I finished talking and ran through the exercises, the hour and a half was finished. And so nobody actually practiced any of the exercises on the first day'* (B2, 96-106)
Programme	Resources/materials	*'The slides didn't work for me this time....You can't lock that room.... once or twice I didn't*

	influenced delivery of SOLAS – education and exercise content	*bring the laptop at all and I just had to print it out, all of the slides on A4 laminate and so we talked all the slides....'* (F1, 207-240) *'I think I only left out maybe three* [exercise] *stations, something like that. Because we didn't have a bouncer and...we didn't have a bed'* (C1, 113-121) *'Nothing but positive feedback for all the content and the-the resources...I just think they complimented the - the education fantastically, I just thought they added much more to the programme than not having these resources.'* (E1, 414-415)

Integrating qualitative and quantitative findings

Each theme was further analysed according to the quantitative data comparing physiotherapists who scored 'high' (i.e. ≥80%; physiotherapists D1, C1, E1 and G1) to those who scored 'moderate' (i.e. ≥50-79%; B1, B2, C2 and F1). Physiotherapist A1 was not included in this analysis as her score was categorised as 'moderate' by direct observations and 'high' by audio-recorded data at 78.7% and 81.3% respectively. A difference between these physiotherapist groups was found in only one theme, '*Group size influenced delivery of SOLAS*'. Physiotherapists who scored higher on the fidelity assessments (average group size of 2.5 participants), believed it was easier to deliver goal-setting as intended with smaller groups. Conversely, physiotherapists with moderate fidelity scores (average group size of 3.7 participants) felt it was harder to facilitate goal-setting as intended with less numbers and believed it would be easier with bigger groups due to better engagement and group discussion. Further details of the triangulation are provided in Table 5 where a meta-matrix was used to compare between findings from both datasets. For the most part, convergence was found between the qualitative and quantitative data, though four qualitative themes relating to influencing factors had no corresponding quantitative data (silence). No areas of discrepancy were found.

Table 5: Triangulation of quantitative and qualitative results – convergence and discrepancy of findings

Outcome of interest	Quantitative fidelity findings		Qualitative interview findings	Convergence/ Discrepancy/ Silence
Fidelity findings:	**Overall fidelity of content scores**	High fidelity (>80%) • Direct Observations: 82.7% • Self-Report: 92.7% • Audio-Recordings: 81.7%	Overall physiotherapists felt that their fidelity was good. Some adaptations and deviations were discussed as occurring in the delivery of the following aspects of the programme: • Goal-setting (Introduction and Review, Review and Planning categories) • Education content • Exercise content • Use of programme materials • Duration of session 1	Convergence
	SOLAS categories scoring below 80% fidelity			
	Materials	Moderate fidelity (50-79%) • Direct Observations: 72.1% • Audio-Recordings: 61.1%		
	Introduction and Review	• Audio-recordings: 76.2%		
	Review and Planning	• Direct Observations: 77.1% • Audio-Recordings: 69.8%		
	Fidelity of duration – sessions significantly different to intended duration			
	Session 1	• Education duration: 58.9'* • Exercise duration: 31.4'*		
Factors influencing	**Correlation between quantitative variables and fidelity scores**		*Theme 1: Physiotherapist knowledge and experience influenced delivery of SOLAS - education content*	Convergence

fidelity findings:	Physiotherapist Experience (years qualified)	• Audio-recordings: -0.333 (p=0.011)***		
	Physiotherapist Group Experience (years delivering group physiotherapy)	• Self-report: -0.430 (p=0.018)***		
	Physiotherapist Post-Training Evaluation Score (%)	• Direct observations: 0.581 (p=0.003)** • Audio-recordings: 0.314 (p=0.018)***	*Theme 1: Physiotherapist knowledge and experience influenced delivery of SOLAS - education content*	Convergence
	Group size (average numbers of participants present)	• Direct observations: -0.434 (p=0.034)***	*Theme 3: Group size influenced delivery of SOLAS – goal-setting, use of materials*	Convergence
	No corresponding quantitative data		*Theme 2: Individual needs influenced delivery of SOLAS - education, exercise, goal-setting*	Silence
			Theme 4: Group dynamics influenced delivery of SOLAS - goal-setting	
			Theme 5: Amount of education content influenced delivery of SOLAS – duration	
			Theme 6: Resources/materials influenced delivery of	

		SOLAS – education and exercise content	
***p<0.001, **p<0.005, ***p<0.05**			

DISCUSSION

The aim of this mixed methods study was to explore and evaluate fidelity of delivery within a feasibility trial of a complex behaviour change intervention using multiple assessment methods. The study found good agreement between researcher-delivered direct observation and audio-recorded fidelity assessment methods, with lower agreement found between provider self-report and researcher-delivered methods. The intervention content was delivered overall with high fidelity, with some variation between physiotherapists and between certain intervention categories. The intervention duration was found to have deviated significantly from intended during the first session only. Subsequently, qualitative interviews with physiotherapists confirmed these fidelity findings. Finally, both qualitative and quantitative data showed that physiotherapists' knowledge and previous experience, as well as the group size, were factors that influenced their fidelity of intervention delivery. The qualitative data contributed further, and postulated additional participant and programme-level factors as aspects that also influenced the overall fidelity results.

Agreement between direct observations and audio-recordings for assessing the fidelity of delivery was found to be excellent. [40] Agreement between both of these methods and provider self-report assessment was lower, as providers consistently rated themselves higher than the independent raters. These findings are perhaps unsurprising, as both direct observations and audio-recordings were rated by the same researcher, and numerous previous studies have shown that providers' subjective assessments of fidelity are often rated higher than independent assessments. [31,32,41] Taking direct observations as the commonly-cited 'gold standard', [11,18,20] these findings reinforce that self-report methods may not be the most accurate method for assessing fidelity in a complex behaviour change study. However, they may still have their place for recording data and also for enhancing fidelity to the protocol by serving as an aide memoire for providers. [42] Although direct observations and audio-recordings have their own limitations [24,43], previous piloting of these assessment methods found that they were feasible and acceptable to physiotherapists. [27] Additionally, the good agreement between audio-recordings and direct observations found in this study suggests that audio-recordings may be a viable alternative with limited resources, as has been done in similar interventions. [44] However, where resources allow, a combination of multiple quantitative methods may provide the most in-depth assessment of fidelity.

One of the key study findings of this study was reinforcing the value of using mixed methods research for the assessment of fidelity. This approach was emphasised in the recent MRC guidelines for conducting process evaluations of complex interventions [6], and is becoming increasingly used in the widespread implementation of evidence-based interventions [45] but does not yet appear to be common practice within fidelity assessments of behaviour change interventions. [10,21] The integration of quantitative and qualitative results enabled the triangulation of findings to provide a better overall picture of the fidelity of the SOLAS intervention and its influencing factors. The importance of the qualitative contribution to answering the 'why' question is evident in the fact that the physiotherapist interviews unearthed strong participant and programme-level factors associated with fidelity results that were not apparent from the quantitative data alone. Whilst this may be predominantly due to the focus of the quantitative analysis on physiotherapist-level variables which were chosen based on existing literature, the participant and programme-level factors identified by this analysis such as group dynamics or amount of programme content may have been difficult to quantitatively analyse to demonstrate association with fidelity results.

This study found that the factors that may influence the fidelity of an interventions' delivery can occur on three levels – provider, participant and programme. Where previous studies have explored factors that have influenced fidelity of intervention delivery, many have focused solely on provider-level factors, demonstrating associations between fidelity and factors such as provider training or skills. [13,19,21,35] The findings of this study have valuable implications for future studies that aim to assess and enhance fidelity of similar interventions as they indicate that planning for fidelity should include considering potential influencing factors at each of these three levels. These results are consistent with recent conclusions by Masterson-Algar et al. in a stroke rehabilitation setting [14] who found that investigating fidelity within clinical trials should also take the individual needs of patients into account, and also concur with the findings of an education-based intervention that found the most common reason for adaptation within intervention delivery was insufficient time. [46]

On the physiotherapist-level, better knowledge of the intervention content and structure was found to positively correlate with quantitative fidelity scores, with a causative link established via the qualitative investigation. This echoes previous findings by Huijg et al. who showed that physiotherapist skill level was one of the most important predictors of fidelity. [21] A more targeted approach to enhancing fidelity in future interventions may

therefore be warranted, such as identifying physiotherapists at higher risk of lower fidelity using post-training evaluation scores and employing more focused fidelity assessment of delivery or further training for these providers, [13] as has been previously employed in similar interventions. [10,44] The results of the study also showed that physiotherapists with more experience of certain aspects tended to emphasise these at the expense of delivering other components as comprehensively as intended in the protocol. These experience-based adaptations invoke the well-established issue of adaptation versus fidelity. For years, research has debated the concept of fidelity versus adaptation, with the case made for both strict fidelity and for modifying interventions. [47] A third view is that both fidelity and adaptation are essential, and achieving an appropriate balance between both can allow an intervention to maximise its effectiveness, while being generalizable and flexible enough to be implementable. [48,49] To achieve this, our fidelity checklists included components that encouraged elements of treatment individualisation (e.g. individualised feedback regarding exercises). However, it may be that these checklists still did not allow for enough individualisation within delivery, an aspect that should be considered by other researchers seeking to conduct similar fidelity assessments.

A limitation of this study was the timing of the interviews, which did not allow a 'pure' convergent/triangulation design. Typically, the qualitative and quantitative methods occur concurrently in this design, [28] however, they were scheduled to take place after physiotherapists had experienced delivery of an entire six-week SOLAS intervention. Although a sequential explanatory design [30] where quantitative results were analysed prior to completing the interviews might have enabled further probing of the factors influencing fidelity, interviews were conducted within two weeks of the intervention completion to minimise recall bias. Due to time constraints it was not possible to have the quantitative data collected and analysed beforehand. Finally, this study mostly focuses on the adherence of delivery (e.g. intervention content and duration) and does not address the *quality* or competence of delivery of SOLAS (e.g. interpersonal or communication style of the physiotherapist), or use of specific BCTs, which is being addressed in a separate publication. This study also does not examine the broader aspects of fidelity such as provider training or participant receipt, as these were beyond the scope of this publication and will be addressed in a future paper.

CONCLUSIONS

In process evaluations and fidelity assessments of large scale complex interventions it is often recommended to complete and report the results of the fidelity assessment before the trial outcomes so as not to bias reporting. [50] Future work will investigate the relationship between this evaluation of fidelity of delivery and the SOLAS feasibility trial outcomes (analysis currently underway), enabling a potentially more insightful and accurate interpretation of findings. This study also has valuable implications for further research and the overall science of fidelity as it contributes much-needed information to the limited current evidence for the application of fidelity assessment methods within the area of complex behaviour change. The findings have demonstrated how multiple quantitative methods can be used to assess the fidelity of delivery of a complex behaviour change intervention, and that a combination of methods may be most suitable, depending on their acceptability and available resources. We have also shown how the use of a mixed methods approach, integrating both quantitative and qualitative data, provides a more insightful understanding of the factors influencing fidelity.

AUTHORS' CONTRIBUTION

ET conceptualized and designed the study and conducted the data collection and analysis. JM and DAH provided guidance on the design. DAH is the principal investigator of the SOLAS trial and obtained funding for this study within the trial. All authors provided methodological consultation and were involved in substantial contributions to reviewing of the drafts. All authors approved the final version of the manuscript.

ACKNOWLEDGEMENTS

This study was funded as part of Health Research Award HRA_HSR/2012/24 by the Health Research Board of Ireland. The funding source had no other involvement in this publication. The authors would like to thank the physiotherapists and participants for taking part in the study. We would also like to thank Alison Keogh for her help in the inter-rater reliability analysis of the audio-recorded data and physiotherapist interviews and Associate Professor Suzanne Guerin for facilitating the interviews.

CONFLICTS OF INTEREST

None

REFERENCES:

1. Gearing RE, El-Bassel N, Ghesquiere A, Baldwin S, Gillies J, Ngeow E. Major ingredients of fidelity: a review and scientific guide to improving quality of intervention research implementation. *Clin. Psychol. Rev.* 2011;31(1):79-88
2. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implementation Science* 2007;2(40):1-9
3. Borrelli B, Sepinwall D, Ernst D, et al. A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *J. Consult. Clin. Psychol.* 2005;73(5):852-60
4. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337(a1655):1-6
5. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348
6. Moore GF, Audrey S, Barker M, et al. *Process evaluation of complex interventions: Medical Research Council guidance*, 2015.
7. Hoffmann TC, Walker MF, Langhorne P, Eames S, Thomas E, Glasziou P. What's in a name? The challenge of describing interventions in systematic reviews: analysis of a random sample of reviews of non-pharmacological stroke interventions. *BMJ Open* 2015;5(11)
8. Schinckus L, Broucke S, Housiaux M. Assessment of implementation fidelity in diabetes self-management education programs: a systematic review. *Patient. Educ. Couns.* 2014;96(1):13-21
9. Toomey E, Currie-Murphy L, Matthews J, Hurley DA. Implementation fidelity of physiotherapist-delivered group education and exercise interventions to promote self-management in people with osteoarthritis and chronic low back pain: A rapid review Part II. *Man. Ther.* 2015;20(2):287-94
10. Mars T, Ellard D, Carnes D, Homer K, Underwood M, Taylor SJC. Fidelity in complex behaviour change interventions: a standardised approach to evaluate intervention integrity. *BMJ Open* 2013;3(11)
11. French SD, Green SE, Francis JJ, et al. Evaluation of the fidelity of an interactive face-to-face educational intervention to improve general practitioner management of back pain. *BMJ Open* 2015;5(7)
12. Perepletchikova F, Hilt LM, Chereji E, Kazdin AE. Barriers to implementing treatment integrity procedures: survey of treatment outcome researchers. *J Consult Clin Psychol* 2009;77(2):212-8
13. Wang B, Stanton B, Deveaux L, et al. Factors influencing implementation dose and fidelity thereof and related student outcomes of an evidence-based national HIV prevention program. *Implement Sci* 2015;10(1):44
14. Masterson-Algar P, Burton CR, Rycroft-Malone J, Sackley CM, Walker MF. Towards a programme theory for fidelity in the evaluation of complex interventions. *Journal of Evaluation in Clinical Practice* 2014;20(4):445-52
15. Langford R, Bonell C, Jones H, Campbell R. Obesity prevention and the Health promoting Schools framework: essential components and barriers to success. *The International Journal of Behavioral Nutrition and Physical Activity* 2015;12:15
16. Mowbray CT, Holter MC, Teague GB, Bybee D. Fidelity Criteria: Development, Measurement, and Validation. *American Journal of Evaluation* 2003;24(3):315-40
17. Poltawski L, Norris M, Dean S. Intervention fidelity: Developing an experience-based model for rehabilitation research. *J. Rehabil. Med.* 2014;46(7):609-15

18. Bellg AJ, Borrelli B, Resnick B, et al. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol.* 2004;23(5):443-51
19. Lorencatto F, West R, Christopherson C, Michie S. Assessing fidelity of delivery of smoking cessation behavioural support in practice. *Implementation Science* 2013;8(40)
20. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J. Public Health Dent.* 2011;71(s1):S52-S63
21. Huijg JM, Dusseldorp E, Gebhardt WA, et al. Factors associated with physical therapists' implementation of physical activity interventions in The Netherlands. *Phys. Ther.* 2015;95(4):539-57
22. McKenna JW, Flower A, Ciullo S. Measuring Fidelity to Improve Intervention Effectiveness. *Intervention in School and Clinic* 2014;50(1):15-21
23. Resnicow K, Davis M, Smith M, et al. How best to measure implementation of school health curricula: a comparison of three measures. *Health Education Research* 1998;13(2):239-50
24. Breitenstein SM, Gross D, Garvey CA, Hill C, Fogg L, Resnick B. Implementation fidelity in community-based interventions. *Res. Nurs. Health* 2010;33(2):164-73
25. Hurley DA, Hall AM, Currie-Murphy L, et al. Theory-driven group-based complex intervention to support self-management of osteoarthritis and low back pain in primary care physiotherapy: protocol for a cluster randomised controlled feasibility trial (SOLAS). *BMJ Open* 2016;6(1):e010728
26. Hurley DA, Murphy LC, Hayes D, et al. Using intervention mapping to develop a theory-driven, group-based complex intervention to support self-management of osteoarthritis and low back pain (SOLAS). *Implementation Science* 2016;11(1):1-29
27. Toomey E, Matthews J, Guerin S, Hurley DA. Development of a Feasible Implementation Fidelity Protocol Within a Complex Physiotherapy-Led Self-Management Intervention. *Phys. Ther.* 2016
28. Creswell JW, Plano Clark VL. *Designing and Conducting Mixed Methods Research*. 2 ed. Thousand Oaks, CA: SAGE Publications, 2011.
29. Barbour RS. The case for combining qualitative and quantitative approaches in health services research. *J. Health Serv. Res. Policy* 1999;4(1):39-43
30. Ivankova NV, Creswell JW, Stick SL. Using Mixed-Methods Sequential Explanatory Design: From Theory to Practice. *Field Methods* 2006;18(1):3-20
31. Hardeman W, Michie S, Fanshawe T, Prevost AT, McLoughlin K, Kinmonth AL. Fidelity of delivery of a physical activity intervention: Predictors and consequences. *Psychol. Health* 2008;23(1):11-24
32. Breitenstein SM, Fogg L, Garvey C, Hill C, Resnick B, Gross D. Measuring Implementation Fidelity in a Community-Based Parenting Intervention. *Nursing research* 2010;59(3):158-65
33. Perepletchikova F, Kazdin AE. Treatment Integrity and Therapeutic Change: Issues and Research Recommendations. *Clinical Psychology: Science and Practice* 2005;12(4):365-83
34. Garbacz L, Brown D, Spee G, Polo A, Budd K. Establishing Treatment Fidelity in Evidence-Based Parent Training Programs for Externalizing Disorders in Children and Adolescents. *Clin Child Fam Psychol Rev* 2014;17(3):230-47
35. Taylor CA, Shaw RL, Dale J, French DP. Enhancing delivery of health behaviour change interventions in primary care: a meta-synthesis of views and experiences of primary care nurses. *Patient Educ. Couns.* 2011;85(2):315-22
36. Braun V, Clarke V. Using thematic analysis in psychology. *Qual. Res. Psychol.* 2006;3(2):77-101
37. Guerin S, Hennessy E. Pupils' definitions of bullying. *Eur J Psychol Educ* 2002;17(3):249-61
38. Wendler MC. Triangulation using a meta-matrix. *J. Adv. Nurs.* 2001;35(4):521-5
39. Farmer T, Robinson K, Elliott SJ, Eyles J. Developing and implementing a triangulation protocol for qualitative health research. *Qualitative Health Research* 2006;16(3):377-94

40. Fleiss J. Reliability of Measurement. In: Fleiss J, ed. The design and analysis of clinical experiments. New York: Wiley, 1986:1-32.
41. Carroll K, Nich C, Rounsaville B. Utility of Therapist Session Checklists to Monitor Delivery of Coping Skills Treatment for Cocaine Abusers. *Psychotherapy Research* 1998;8(3):307-20
42. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Publ Health Dentistry* 2011;71:S52 - S63
43. Leventhal H, Friedman MA. Does establishing fidelity of treatment help in understanding treatment efficacy? Comment on Bellg et al. (2004). *Health Psychol.* 2004;23(5):452-6
44. Pincus T, Anwar S, McCracken LM, et al. Delivering an Optimised Behavioural Intervention (OBI) to people with low back pain with high psychological risk; results and lessons learnt from a feasibility randomised controlled trial of Contextual Cognitive Behavioural Therapy (CCBT) vs. Physiotherapy. *BMC Musculoskelet. Disord.* 2015;16:147
45. Palinkas LA, Aarons GA, Horwitz S, Chamberlain P, Hurlburt M, Landsverk J. Mixed Method Designs in Implementation Research. *Administration and Policy in Mental Health* 2011;38(1):44-53
46. Hill LG, Maucione K, Hood BK. A focused approach to assessing program fidelity. *Prev Sci* 2007;8(1):25-34
47. Dane AV, Schneider BH. Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin. Psychol. Rev.* 1998;18(1):23-45
48. Castro FG, Barrera M, Jr., Martinez CR, Jr. The cultural adaptation of prevention interventions: resolving tensions between fidelity and fit. *Prev Sci* 2004;5(1):41-5
49. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology* 2008;41(3-4):327-50
50. Oakley A, Strange V, Bonell C, Allen E, Stephenson J. Health services research: process evaluation in randomised controlled trials of complex interventions. *BMJ* 2006;332:413

Supplementary Files

Supplementary File 1: Quantitative fidelity checklists used

INTERVENTION SESSION CHECKLIST (OBSERVATION/AUDIO-RECORDING):

Cover sheet (completed for each session)

Date:	
Venue:	
Physiotherapist Name:	
Other staff involved:	Name(s): Role: (e.g. set-up/delivery/support)
Session number (tick):	1☐ 2☐ 3☐ 4☐ 5☐ 6☐
Start time (class):	
Finish time (class):	
Time spent on education (mins):	
Time spent on exercise (mins):	
Method of Observation:	In-vivo/Audio-recording
Adverse event(s)/issue(s) (circle): **Y/N**	If yes give brief details:

Deviations from protocol/proscribed components delivered? **General notes on fidelity of session:**	Yes/No Details:				
INTERVENTION COMPONENT CHECKLIST:		**YES (2)**	**NO (0)**	**ATTEMPTED (1)**	**N/A**
Session 1:					
Materials					
	Intervention folder given to participants	☐	☐	☐	
	Use of **participant intervention folder** actively facilitated **throughout** session	☐	☐	☐	
	Name stickers/badges given to participants	☐	☐	☐	
	Pens offered/provided	☐	☐	☐	
	Powerpoint slides used	☐	☐	☐	
Introduction					
	Introductions/welcome made	☐	☐	☐	
	Set **clear expectations - aims**, **content** and **structure** of programme addressed	☐	☐	☐	
	Rationale for **weekly attendance** provided	☐	☐	☐	
Education					

Rationale for **self-management** provided	☐	☐	☐	
Posture addressed	☐	☐	☐	
Cycle of change addressed	☐	☐	☐	
Disease characteristics, prevalence and causes of OA/CLBP addressed	☐	☐	☐	
Recommended activity levels/benefits of exercise addressed	☐	☐	☐	
Individual reflection on activity/recommendations facilitated	☐	☐	☐	
Goal setting introduced	☐	☐	☐	
Action planning introduced	☐	☐	☐	
Attention drawn to **Action Plan sheets** within intervention folder, **encouraged use**	☐	☐	☐	
Attention drawn to **Activity Diaries** within intervention folder, **encouraged use**	☐	☐	☐	
Participants given a chance/encouraged to **contribute to discussion**	☐	☐	☐	
Exercise				
Room set up for exercise (equipment, sheets)	☐	☐	☐	
Protocol exercises **demonstrated**	☐	☐	☐	
Attention drawn to **Exercise Programme Diary** within intervention folder, **encouraged use**	☐	☐	☐	
Rationale for **exercises** provided	☐	☐	☐	
Participants given a chance to **attempt and practice protocol** exercises	☐	☐	☐	
Individual feedback provided	☐	☐	☐	☐
Review and Planning				
Session review - **activity levels and goal setting recap**	☐	☐	☐	
Proscribed components delivered?	☐(-2)	☐		
Total score (Yes = 2, Attempted = 1, No =0)				
Overall Adherence score				

Session 2:

INTERVENTION COMPONENT CHECKLIST:	**YES (2)**	**NO (0)**	**ATTEMPTED (1)**	**N/A**

Materials					
Use of **participant intervention folder** actively facilitated **throughout** session	☐	☐	☐		
Name stickers/badges given to participants	☐	☐	☐		☐
Pens offered/provided	☐	☐	☐		
Pedometers offered	☐	☐	☐		
Powerpoint slides used	☐	☐	☐		
Recap and Review					
Welcome back made	☐	☐	☐		
Review of previous week **activities/action plans**	☐	☐	☐		
Education					
Set clear expectations – content of session outlined	☐	☐	☐		
Activity-rest cycle and **pacing** explained	☐	☐	☐		
Individual reflection on **current pacing/activity-rest** facilitated	☐	☐	☐		
Factors influencing pain addressed	☐	☐	☐		
Goal setting facilitated	☐	☐	☐		
Action planning facilitated	☐	☐	☐		
Action Plan sheets use encouraged/facilitated	☐	☐	☐		
Attention drawn to **Walking/Activity Diary** within intervention folder, **encouraged use**	☐	☐	☐		
Participants given a chance/encouraged to **contribute to discussion**	☐	☐	☐		
Exercise					
Room set up for exercise (equipment, sheets)	☐	☐	☐		
Protocol exercises **demonstrated**	☐	☐	☐		☐
Exercise Programme Diary use encouraged	☐	☐	☐		
Rationale for **exercises** provided	☐	☐	☐		
Participants given a chance to **attempt and practice protocol** exercises	☐	☐	☐		
Individual feedback provided	☐	☐	☐		☐
Session Planning and Review					
Session review - **goal setting and action planning recap**	☐	☐	☐		
Proscribed components delivered?	☐(-2)	☐			
Total score (Yes = 2, Attempted = 1, No =0)					

Overall Adherence score				
Session 3:				
INTERVENTION COMPONENT CHECKLIST:	**YES (2)**	**NO (0)**	**ATTEMPTED (1)**	**N/A**
Materials				
Use of **participant intervention folder** actively facilitated **throughout** session	☐	☐	☐	
Tape measures offered	☐	☐	☐	
Pens offered/provided	☐	☐	☐	
Powerpoint slides used	☐	☐	☐	
Recap and Review				
Previous week **Activity Action Plan** reviewed	☐	☐	☐	
Problem-solving of previous week **Activity Action Plan**	☐	☐	☐	☐
Action Plan sheets use encouraged/facilitated	☐	☐	☐	
Education				
Set clear expectations – content of session outlined	☐	☐	☐	
Obesity and effect on pain addressed	☐	☐	☐	
Balance between weight/activity addressed	☐	☐	☐	
Individual reflection on **weight/activity** balance facilitated	☐	☐	☐	
Skills for maintaining healthy weight addressed (e.g. waist measurement, food diary)	☐	☐	☐	
Attention drawn to **Healthy Eating Booklet** within intervention folder, **encouraged use**				
Attention drawn to **Walking/ Activity Diary** within intervention folder, **encouraged use**	☐	☐	☐	
Attention drawn to **Food & Exercise Diary** within intervention folder, **encouraged use**	☐	☐	☐	
Participants given a chance/encouraged to **contribute to discussion**	☐	☐	☐	
Exercise				
Room set up for exercise (equipment, sheets)	☐	☐	☐	
Protocol exercises **demonstrated**	☐	☐	☐	☐

Exercise Programme Diary use encouraged	☐	☐	☐	
Rationale for **exercises** provided	☐	☐	☐	
Participants given a chance to **attempt and practice protocol** exercises	☐	☐	☐	
Individual feedback provided	☐	☐	☐	☐
Session Planning and Review				
Session review - **goal setting and action planning** recap integrating **food and exercise diary**	☐	☐	☐	
Proscribed components delivered?	☐(-2)	☐		
Total score (Yes = 2, Attempted = 1, No =0)				
Overall Adherence score				

Session 4:

INTERVENTION COMPONENT CHECKLIST:	**YES (2)**	**NO (0)**	**ATTEMPTED (1)**	**N/A**
Materials				
Use of **participant intervention folder** actively facilitated **throughout** session	☐	☐	☐	
Pens offered/provided	☐	☐	☐	
Powerpoint slides used	☐	☐	☐	
Midway Recap and Review				
Previous week **Activity Action Plan** reviewed	☐	☐	☐	
Problem-solving of previous week **Activity Action Plan**	☐	☐	☐	☐
Action Plan sheets use encouraged/facilitated	☐	☐	☐	
Walking/ Activity Diary use encouraged	☐	☐	☐	
Education				
Set clear expectations – content of session outlined	☐	☐	☐	
Rationale for using **pain relief** given (e.g. **pain pathway** explained)	☐	☐	☐	
Methods of pain relief addressed (e.g. medication, heat/ice, TENS/acupuncture)	☐	☐	☐	
Individual reflection on use of **pain relief** facilitated	☐	☐	☐	

Participants given a chance/encouraged to **contribute to discussion**	☐	☐	☐	
Exercise				
Room set up for exercise (equipment, sheets)	☐	☐	☐	
Protocol exercises **demonstrated**	☐	☐	☐	☐
Exercise Programme Diary use encouraged	☐	☐	☐	
Rationale for **exercises** provided	☐	☐	☐	
Participants given a chance to **attempt and practice protocol** exercises	☐	☐	☐	
Individual feedback provided	☐	☐	☐	☐
Session Planning and Review				
Session review - **goal setting and action planning** recap	☐	☐	☐	
Proscribed components delivered?	☐(-2)	☐		
Total score (Yes = 2, Attempted = 1, No =0)				
Overall Adherence score				

Session 5:

INTERVENTION COMPONENT CHECKLIST:	**YES (2)**	**NO (0)**	**ATTEMPTED (1)**	**N/A**
Materials				
Use of **participant intervention folder** actively facilitated **throughout** session	☐	☐	☐	
Pens offered/provided	☐	☐	☐	
Relaxation CD offered	☐	☐	☐	
Powerpoint slides used	☐	☐	☐	
Recap and Review				
Previous week **Activity Action Plan** reviewed	☐	☐	☐	
Problem-solving of previous week **Activity Action Plan**	☐	☐	☐	☐
Action Plan sheets use encouraged/facilitated	☐	☐	☐	
Walking/ Activity Diary use encouraged	☐	☐	☐	
Education				
Set clear expectations – content of session outlined	☐	☐	☐	

Recognising and managing flare-ups addressed	☐	☐	☐	
Individual reflection about **flare-ups** facilitated	☐	☐	☐	
Effect of mood on pain addressed	☐	☐	☐	
Participants given a chance/encouraged to **contribute to discussion**	☐	☐	☐	
Exercise				
Room set up for exercise (equipment, sheets)	☐	☐	☐	
Protocol exercises **demonstrated**	☐	☐	☐	☐
Exercise Programme Diary use encouraged	☐	☐	☐	
Rationale for **exercises** provided	☐	☐	☐	
Participants given a chance to **attempt and practice protocol** exercises	☐	☐	☐	
Individual feedback provided	☐	☐	☐	☐
Relaxation Session				
Relaxation techniques explained and **practiced**	☐	☐	☐	
Session Planning and Review				
Session review - **goal setting and action planning** recap with **integration** of **relaxation techniques**	☐	☐	☐	
Proscribed components delivered?	☐(-2)	☐		
Total score (Yes = 2, Attempted = 1, No =0)				
Overall Adherence score				

Session 6:

INTERVENTION COMPONENT CHECKLIST:	**YES (2)**	**NO (0)**	**ATTEMPTED (1)**	**N/A**
Materials				
Use of **participant intervention folder** actively facilitated **throughout** session	☐	☐	☐	
Pens offered/provided	☐	☐	☐	
Handouts on **local resources and supports** provided	☐	☐	☐	
Powerpoint slides used	☐	☐	☐	
Recap and Review				
Previous week **Activity Action Plan** reviewed	☐	☐	☐	
Problem-solving of previous week **Activity Action Plan**	☐	☐	☐	☐

Action Plan sheets use encouraged/facilitated	☐	☐	☐	
Walking/ Activity Diary use encouraged	☐	☐	☐	
Education				
Set clear expectations – content of session outlined	☐	☐	☐	
Core skills of programme reviewed	☐	☐	☐	
Aims of long-term self-management addressed	☐	☐	☐	
Local resources and supports discussed	☐	☐	☐	
Participants given a chance/encouraged to **contribute to discussion**	☐	☐	☐	
Exercise				
Room set up for exercise (equipment, sheets)	☐	☐	☐	
Protocol exercises **demonstrated**	☐	☐	☐	☐
Exercise Programme Diary use encouraged	☐	☐	☐	
Rationale for **exercises** provided	☐	☐	☐	
Participants given a chance to **attempt and practice protocol** exercises	☐	☐	☐	
Individual feedback provided	☐	☐	☐	☐
Session Planning and Review				
Session review – **long term goal setting and action planning** recap	☐	☐	☐	
Activity diaries **use recorded** (if willing)	☐	☐	☐	
Proscribed components delivered?	☐(-2)	☐		
Total score (Yes = 2, Attempted = 1, No =0)				
Overall Adherence score				

Yes = Fully addressed by the Physiotherapist – could not do more

Attempted = Reasonable attempt made to address this – could do more

No = No attempt made to address this

INTERVENTION SESSION CHECKLIST (SELF-REPORT)

Self-report checklists covered similar components to observation and audio as provided above - different cover sheet (completed for each session) provided below

PCCC Site:	Date:	Class: circle 1 2 3 4 5 6	Physiotherapist Name:

PERSONNEL

Other staff involved in setting up class:	Yes ☐ No ☐	Names/Staff Grade:
Other staff involved in providing class:	Yes ☐ No ☐	Names/Staff Grade:

ATTENDANCE

	Present	Absent
Number of Clients:		
Names of non-attenders:	Reasons for non-attendance [if known]	
1. 2. 3. 4.		

CLASS PREPARATION

Time to review materials [mins]:	Time to set up class [mins]:	Time to take down class [mins]:
Start Time:	End Time:	Comments:

CLASS DELIVERY

	Education	Exercise
Time to deliver [mins]:		
Comments:		
Deviations from protocol: Content/time	Yes ☐ No ☐	Yes ☐ No ☐
If 'yes' give details and reason[s]		

TREATMENT-RELATED EVENT – any unforeseen event/issue should be reported to Deirdre Hurley

During class: Yes ☐ No ☐	**After class:** Yes ☐ No ☐	**Reported to Deirdre** Yes ☐ No ☐

When patient is discharged please give this completed form to UCD Research Physiotherapist or scan and email to physiostudy@ucd.ie

Supplementary File 2: Physiotherapist baseline characteristics

	Wave 1				Wave 2					
Physiotherapy Site	A	B	C	D	A	B	C	E	F	G
Physiotherapist	A1	B1	C1	D1	A1	B2	C2	E1	F1	G1
Gender	F	F	F	M	F	F	F	F	F	F
Experience (years qualified)	5	25	8	12	5	25	10	6	19	11
Group experience (years delivering group physiotherapy)	1	3	8	2	1	10	3	1	15	7
Post-training evaluation score (%)	74%	81.5%	90.5%	83.4%	74%	88.5%	81.8%	88.2%	78.1%	94.4%
Previous training in similar interventions (Y/N)	N	Y	Y	Y	N	Y	Y	Y	Y	Y
F= female, M=male, Y=yes, N=no										

Supplementary File 3: Physiotherapist interviews topic guide

SOLAS Programme Semi-Structured Interviews with Physiotherapists

Interview Schedule

Ice-breaker – describe your experience in running group-based programmes for these or other populations before this study. How many weeks of the SOLAS programme did you deliver?

Therapist Views on Experience of Delivering the Programme

- What are your overall impressions of the programme having delivered it for first/second time?
    - Content overall and week by week –education and exercise components – time for each- managing group dynamics – mixed ability/diagnosis and ages of clients
- What aspects did you find easy/challenging to deliver? [content/communication/behaviour change/goal setting/action planning – these were emphasised in training]
- Views on feasibility of class size of up to 8 clients with one physiotherapist to deliver? [none delivered a class of this size – their views on running small numbers v the target of 8]
- How well do you feel you delivered the programme as intended from the training received?
    - Following the slide content/script – was it difficult/did you want to edit - which parts?
    - Content/needs supportive delivery style/ behaviour change techniques
    - Views on giving advice/setting goals with patients and following through and being needs supportive (using SDT) e.g. using non-controlling language, enabling patient input and choice; providing positive and personalised feedback to patients?
    - Difficult/constrained by research?
- For the aspects not delivered as planned from training give reasons –
    - prompt on potential barriers to delivery [the availiability of resources (e.g. staffing, suitable venue, administrative staff, time constraints); appropriate patient selection and screening, patient uptake and engagement with programme and the potential need to individualise treatment within group]
- How much additional work did delivering the programme and participating in this study place upon you? [Specify – preparation time – reading the manual and supplementary materials, setting up the venue, time to deliver – on top of other work, completion of treatment record forms after each class, completion of post training questionnaires]?
    - Is this acceptable? What modifications would you suggest for future waves?
- What are your impressions of the resources provided to you to support delivery of the programme? [Training Manual /Intervention Manual/Intervention Slides/SOLAS poster]
    - How much have you used them? What was useful/not useful in terms of helping you deliver the programme as intended?
    - Views on continuing to use powerpoint versus flip chart or handouts only?
    - Suggestions for modifications for future waves
- What are your impressions of the venue in terms of its suitability for delivering the programme? [Prompt – accessibility, space, equipment for delivering education and exercise component number of stations, sound quality to allow communication to the

group/individuals during exercises

Views on sustainability of the Programme

- What is your opinion of the feasibility of delivering the programme in the future within the study/within normal practice?
- Would you deliver the programme outside the study?
- How would you deliver it? What materials would you use? What would you leave out? Who would you deliver it to?

Views on Research Elements of the Programme

- Views on research elements of the programme [pre and post training questionnaires, direct observation during classes, audio recordings, treatment record forms]
    - Intrusiveness/time/feasibility Any suggestions for modifications for future waves
- Views on level and modes of communication with the research team throughout the study from training to completion of this wave
    - Suggestions for modifications for future waves

Supplementary File 4: Agreement between assessment methods				
	Direct Observation v Self-Report (n=24)	Direct Observation v Audio-Recordings (n=24)	Audio-Recordings v Self-Report (n=60)	Audio-Recordings Rater 1 v Rater 2 (Inter-rater reliability) (n=12)
Overall agreement	74.6%	86.6%	75.4%	84.6%
% Agreement per category:				
Materials	74.5%	82.67%	70.1%	84.6%
Introduction and Review	65.3%	86.5%	57%	81.6%
Education	84.6%	90.3%	87.3%	76.7%
Exercise	70.8%	86%	78%	83.3%
Review and Planning	50%	76.2%	46%	100%

BMJ Open

Using mixed methods to assess fidelity of delivery and its influencing factors in a complex self-management intervention for people with osteoarthritis and low back pain

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-015452.R2
Article Type:	Research
Date Submitted by the Author:	31-Mar-2017
Complete List of Authors:	Toomey, Elaine; National University of Ireland - Galway, School of Psychology Matthews, James; University College Dublin, Hurley, Deirdre; University College Dublin, Physiotherapy and Performance Science
Primary Subject Heading:	Sports and exercise medicine
Secondary Subject Heading:	Sports and exercise medicine, Rheumatology, Qualitative research, Patient-centred medicine
Keywords:	implementation fidelity, PRIMARY CARE, complex interventions, behaviour change

Title

Using mixed methods to assess fidelity of delivery and its influencing factors in a complex self-management intervention for people with osteoarthritis and low back pain

Authors

Elaine Toomey[a]*, James Matthews[b], Deirdre A. Hurley[b].

[a]School of Psychology, Arts Millennium Building, National University of Ireland, Galway, Ireland

[b]School of Public Health, Physiotherapy and Sports Science, Health Sciences Centre, University College Dublin, Belfield, Dublin 4, Ireland

*Corresponding author

Corresponding author contact details

Email: elaine.toomey@nuigalway.ie

Phone: 00353-91494458

Word count

4229

ABSTRACT:

Objectives and design: Despite an increasing awareness of the importance of fidelity of delivery within complex behaviour change interventions, it is often poorly assessed. This mixed methods study aimed to establish the fidelity of delivery of a complex self-management intervention and explore the reasons for these findings using a convergent/triangulation design.

Setting: Feasibility trial of the Self-management of Osteoarthritis and Low back pain through Activity and Skills (SOLAS) intervention, delivered in primary care physiotherapy.

Methods and outcomes: 60 SOLAS sessions were delivered across seven sites by nine physiotherapists. Fidelity of delivery of pre-specified intervention components was evaluated using 1) audio-recordings (n=60), direct-observations (n=24) and self-report checklists (n=60) and 2) individual interviews with physiotherapists (n=9). Quantitatively, fidelity scores were calculated using percentage means and standard deviations of components delivered. Associations between fidelity scores and physiotherapist variables were analysed using Spearman's correlations. Interviews were analysed using thematic analysis to explore potential reasons for fidelity scores. Integration of quantitative and qualitative data occurred at an interpretation level using triangulation.

Results: Quantitatively, fidelity scores were high for all assessment methods; with self-report (92.7%) consistently higher than direct-observations (82.7%), or audio-recordings (81.7%). There was significant variation between physiotherapists' individual scores (69.8% - 100%). Both qualitative and quantitative data (from physiotherapist variables) found that physiotherapists' knowledge (Spearman's association at p=0.003) and previous experience (p=0.008) were factors that influenced their fidelity. The qualitative data also postulated participant-level (e.g. individual needs) and programme-level factors (e.g. resources) as additional elements that influenced fidelity.

Conclusion: The intervention was delivered with high fidelity. This study contributes to the limited evidence regarding fidelity assessment methods within complex behaviour change interventions. The findings suggest a combination of quantitative methods is suitable for the

assessment of fidelity of delivery. A mixed methods approach provided a more insightful understanding of fidelity and its influencing factors.

Article summary – strengths and limitations

- This mixed methods investigation of fidelity of delivery and its influencing factors provides valuable information on fidelity assessment methods and factors to be considered in developing and evaluating complex behaviour change interventions
- The novel use of mixed methods to assess fidelity in this study enabled increased certainty in findings where qualitative data corroborated the quantitative results
- This study does not explore the fidelity of the *quality* of delivery (e.g. therapist competence) or specific behaviour change techniques (BCTs) which will be reported in a separate publication

USING MIXED METHODS TO ASSESS FIDELITY OF DELIVERY AND ITS INFLUENCING FACTORS IN A COMPLEX SELF-MANAGEMENT INTERVENTION

INTRODUCTION

Fidelity of delivery is an important, yet often overlooked aspect of behaviour change interventions and has been defined as the degree to which an intervention, treatment or program is delivered as intended by the intervention developers. [1,2] Without adequately addressing fidelity in behaviour change research, it is uncertain that changes observed in study outcomes are due to the influence of the independent variable (the intervention being investigated) and not due to variability in its implementation, e.g. extraneous elements that may have been added (either accidentally or purposely), or essential elements of the intervention that were omitted. [3] In particular, as behaviour change interventions are often complex interventions that typically involve several components with the potential to affect or influence outcomes separately, it is especially important to incorporate adequate fidelity planning and assessment into the development of interventions of this nature. [4]

Despite a recent increased emphasis on improved assessment and reporting of what happens within complex behaviour change interventions, [5,6] fidelity is still poorly addressed within this context, with few examples of fidelity being assessed comprehensively or reported adequately. [7-9] Where studies have assessed fidelity within a behaviour change healthcare context, there is often limited exploration of the factors that might have influenced that fidelity. [10,11] Previous work that *has* specifically examined influencing factors in areas of public health, obesity and stroke research found that provider-level variables, such as experience, knowledge or skills, may influence fidelity of delivery. [12-15] Although the use of both quantitative and qualitative methods has been previously recommended to comprehensively assess fidelity, [16-18] this guidance is not consistently followed. Consequently, the use of quantitative methods in isolation may not allow for exploration of the factors influencing fidelity, knowledge of which could improve how fidelity is enhanced and assessed in future similar interventions. [19] For example, French et al. [11] used audio-recordings and self-report methods to assess the fidelity of delivery of an educational

intervention for general practitioners. The authors acknowledged that the quantitative study design did not allow them to explore the reasons for variations in fidelity scores found.

According to existing guidelines for addressing fidelity within behaviour change research developed by the National Institutes of Health Behaviour Change Consortium (NIHBCC), [18,20] the fidelity of delivery of the intervention by providers is one particularly important aspect of fidelity. This aspect considers strategies that enhance the fidelity of delivery (e.g. using treatment manuals or intervention protocols) and methods that assess this delivery (e.g. provider self-report, audio or video-recorded observations and direct *in vivo* observations). However, although previous research has advocated a combination of these strategies in order to assess fidelity in-depth, [21,22] limited examples exist within the literature. Additionally, few studies have explored the relationship between these methods, and the accuracy of potentially more feasible methods against the 'gold standard' of direct observations using pre-specified criteria [23] has been poorly investigated. [18,19,24] As a result, there is little evidence to justify the selection of one method over another, or to inform the use of multiple methods simultaneously.

The current study is set within the context of the SOLAS (Self-Management of Osteoarthritis and Low back pain through Activities and Skills) feasibility trial [ISRCTN49875385]. [25] The trial aims to evaluate the feasibility of providing the SOLAS intervention (experimental group) to promote self-management for patients with osteoarthritis (OA) of the hip/ knee and/or chronic low back pain (CLBP) compared to usual physiotherapy, which will serve as the pragmatic control group in order to determine the feasibility of moving to a full scale trial by following the MRC guidelines. The intervention consists of six weekly sessions of 90 minutes to be delivered by a primary care physiotherapist to a group of six to eight people. Each session is divided into education and exercise sections (each approximately 45 minutes in duration). During the exercise section, participants are provided with an opportunity to attempt and practice pre-specified exercises, while the education section is further split into four categories: Materials - participants are provided with materials intended to supplement and enhance understanding and uptake of skills; Introduction and Review - at the start of each session the physiotherapist introduces session aims and reviews goals and action plans with participants; Education - the physiotherapist facilitates a group discussion on targeted self-management skill or behaviour of the session using Powerpoint slides and projector; Review and Planning - before the session concludes, the physiotherapist recaps participants' planned

activity levels and action plans. For each category, a number of components (i.e. specific elements or activities which varied according to the session) were intended to be delivered or addressed during the session by the physiotherapist (e.g. provision of pedometers, provision of information on balanced weight and healthy eating), as detailed in an intervention manual [25,26] (summarised in Table 1). Prior to participation, physiotherapists were provided with the manual during a two-day training course where background variables were collected, in addition to a post-training assessment. [26] Details of the development and testing of the fidelity protocol used in this study are described elsewhere. [27] This study also does not explore the fidelity of the *quality* of delivery (e.g. therapist competence) or specific behaviour change techniques (BCTs) which will be reported elsewhere.

The aim of this study was to establish the fidelity of delivery of a complex behaviour change intervention and the reasons for these findings using a mixed methods approach. Specifically, the study objectives were:

1) To evaluate the agreement of multiple methods for assessing fidelity of delivery
2) To establish the fidelity of delivery of the SOLAS complex behaviour change intervention
3) To explore the potential factors that may have influenced these fidelity results

Table 1: Structure, content and samples of SOLAS intervention

Session number	Self-management behaviours/skills targeted	Session section (intended duration)	Session category and component example*
1	• Physical activity • Goal-setting	Education (45 mins)	• Materials o *Provide participant handbook* • Introduction and Review o *Introduce programme aims* • Education o *Provide information on benefits of exercise/physical activity* • Review and Planning o *Review of SOLAS programme and activity levels*
		Exercise (45 mins)	• Exercise o *Demonstrate protocol exercises*
2	• Pacing (balancing activity/rest) • Physical activity • Goal-setting	Education (45 mins)	• Materials o *Offer pedometers* • Introduction and Review o *Review of previous week activities* • Education o *Provide information on activity-rest cycle and pacing* • Review and Planning o *Review of session and goal-setting*

		Exercise (45 mins)	• Exercise ○ *Provide opportunity for participants to practice exercises*

3	• Balanced weight/ healthy diet
	• Physical activity
	• Goal-setting
		Education (45 mins)	• Materials
			  ○ *Offer tape measures*
			• Introduction and Review
			  ○ *Review of previous session action plan*
			• Education
			  ○ *Provide information on balanced weight and healthy eating*
			• Review and Planning
			  ○ *Review of session and goal-setting*

		Exercise (45 mins)	• Exercise ○ *Provide opportunity for participants to practice exercises*

4	• Managing pain with pain relief
	• Physical activity
	• Goal-setting
		Education (45 mins)	• Materials
			  ○ N/a
			• Introduction and Review
			  ○ *Mid-way review*
			• Education
			  ○ *Provide information on pain management with ice/heat*
			• Review and Planning
			  ○ *Review of session and goal-setting*

		Exercise (45 mins)	• Exercise

			○ *Provide opportunity for participants to practice exercises*
5	• Managing mood (with relaxation) • Physical activity • Goal-setting	Education (45 mins)	• Materials ○ *Offer relaxation CD* • Introduction and Review ○ *Review of previous session action plan* • Education ○ *Provide information on recognising and managing flare-ups* • Review and Planning ○ *Review of session and goal-setting*
		Exercise (45 mins)	• Exercise ○ *Provide opportunity for participants to practice exercises*
6	• Long-term management • Physical activity • Goal-setting	Education (45 mins)	• Materials ○ *Offer 'local resources to support physical activity' leaflet* • Introduction and Review ○ *Review of previous session action plan* • Education ○ *Provide information on local resources and supports* • Review and Planning ○ *Long-term goal setting and action planning*
		Exercise (45 mins)	• Exercise ○ *Provide opportunity for participants to practice exercises*

Samples involved											
	Wave 1				**Wave 2**						**Total n=**
Physiotherapy Site code	A	B	C	D	A	B	C	E	F	G	**10 (7)****
Physiotherapist code	A1	B1	C1	D1	A1	B2	C2	E1	F1	G1	**9*****
Number of participants recruited	4	6	4	2	4	4	5	4	7	5	**45**
Sessions delivered	6	6	6	6	6	6	6	6	6	6	**60**
Direct Observations	3	3	3	3	2	2	3	2	1	2	**24 (40%)**
Self-Report	6	6	6	6	6	6	6	6	6	6	**60 (100%)**
Audio-Recordings (rater 1)	6	6	6	6	6	6	6	6	6	6	**60 (100%)**
Audio-Recordings (rater 2)	3	3	3	3	0	0	0	0	0	0	**12 (20%)**
*A complete list of category components is provided in the Supplementary Files **Seven sites were involved in total, but three of these delivered the intervention in both waves ***One physiotherapist delivered the intervention in both waves											

METHODS

Ethics

Ethical approval for this study was granted by University College Dublin's Human Research Ethics Committee (LS-13-54). Written informed consent was obtained from all physiotherapists and participants prior to any observations or interviews.

Design

This observational study was a convergent/triangulation mixed methods design. [28] This mixed methods approach was chosen as it was felt that thorough integration of findings from both quantitative and qualitative methods would achieve a more comprehensive answer to the study questions by enabling the methods to be 'greater than just the sum of their parts'. [29] Figure 1 graphically illustrates the study design, outlining the sequence of research activities, the priority of the methods and the stage at which integration occurred. [30]

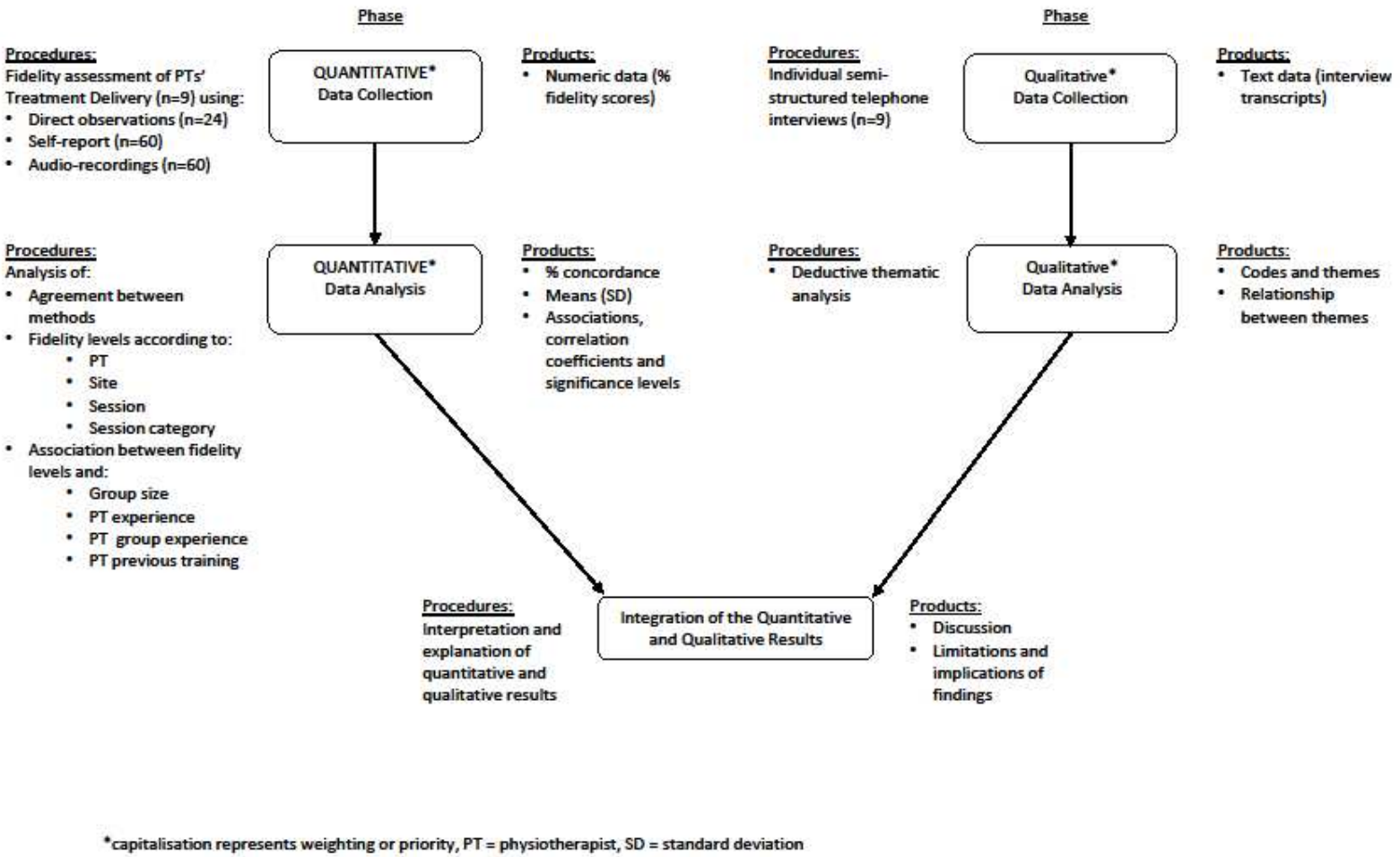

Figure 1: Diagrammatic representation of mixed methods convergent/triangulation design

Quantitative phase

Study sample and procedure:

Data were collected during two of the three waves of the SOLAS feasibility trial (Table 1), representing 71% of the overall trial data. Fidelity of delivery in this study refers to the assessment of both the delivery of session content, i.e. providers deliver the session categories and components as intended (summarised in Table 1) (fidelity of content) and session duration, i.e. providers deliver the session as long as intended (fidelity of duration). Following pilot testing, it was decided to conduct 24 (40% of sessions) randomly selected direct observations (rated by ET), 60 (100%) self-report (rated by the physiotherapists) and 60 (100%) audio-recordings (rated by ET) to assess fidelity of delivery using a-priori checklists (Supplementary File 1) that had been previously found to be feasible for use. [27] To assess inter-rater reliability, 12 sessions (20%) were rated by a second independent rater (AK). [10] Checklists consisted of approximately 25 components for each session, structured according to the SOLAS categories as detailed in Table 1. Components for each session were chosen to address each element specified in the SOLAS intervention manual (summarised in Table 1) [25,26] to be delivered during that session. Each component was rated as 'Yes/Present' equating to a score of two points, 'No/Absent' (zero points), or 'Attempted' (one point). Session duration was documented by all methods, and attendance was recorded by self-report.

Data analysis:

Fidelity data analysis was consistent with standard procedures [19,31,32] using SPSS v20. Specifically, levels of agreement between methods and inter-rater reliability of audio-recorded data were assessed using percentage concordance. Overall mean fidelity of content scores (i.e. percentage of manual-specified components delivered as intended) and fidelity scores according to physiotherapy site, physiotherapist, session and session category were obtained by calculating total actual scores as a percentage of the total possible score. Means data were compared using ANOVAs and Kruskal-Wallis tests. Fidelity of duration was established by calculating the difference between the actual and the intended session duration using a one-sample Wilcoxon test. Levels of fidelity were interpreted as previously reported in the literature, with 80-100% adherence interpreted as 'high' fidelity, 51-79% as 'moderate' and 0-50% as 'low' fidelity. [3,33,34] Finally, the relationship between fidelity scores and 1) the number of participants present (group size) and 2) physiotherapist variables, i.e. experience (years qualified), group experience (years delivering group physiotherapy), knowledge of

intervention (post-training evaluation score) and previous relevant training (Supplementary File 2), were calculated using Spearman's correlation coefficient and Mann-Whitney U test. These physiotherapist variables were chosen for reasons described in the introduction.

Qualitative phase

The aim of the qualitative phase was to explore physiotherapists' opinions of fidelity of intervention delivery and the factors that they felt may have influenced their fidelity. Individual semi-structured telephone interviews were conducted by an experienced qualitative researcher (SG) with each physiotherapist (n=9) within two weeks of intervention delivery completion. A topic guide with specific questions and probes related to fidelity was developed by the corresponding author (ET) (Supplementary File 3). Interviews were audio-recorded and transcribed verbatim. Deductive thematic analysis was used to analyse the interviews as it is a flexible method that works with a range of research questions, including understanding people's experiences of programmes and health-care interventions. [35]

Meaningful units of text were highlighted within each interview, then summarised and coded. Codes dealing with similar issues were grouped across all interviews and refined into themes. The reliability of themes was established by a second reviewer (AK) who independently analysed a randomly selected sample of 50% of the transcript extracts using the coding framework. Percentage agreement was determined between the reviewers' respective coding of extracts. If agreement was less than 70%, consensus on conflicting decisions was obtained through discussion. [36]

Integration

Integration of quantitative and qualitative data occurred at an interpretation level using triangulation methodology. Specifically, a meta-matrix was created to facilitate comparison of the findings (Table 5). [37] This involved presenting the quantitative data in a tabular format alongside summarised qualitative themes, which enabled a transparent approach to determining convergence, discrepancy or silence across the findings of the data sets. [38] Convergence was defined as general agreement between the data sets on the element of comparison (e.g. overall quantitative fidelity score compared to the majority of physiotherapist opinions of their fidelity levels), while discrepancy was defined as general disagreement between the data sets on the element of comparison. [38] Silence was defined as

where one set of results addressed a theme or example, but the other set of results did not yield any relevant data. [38]

RESULTS

Quantitative findings

Agreement:

Agreement between direct observations and self-report was 74.6%, 75.4% between self-report and audio-recordings (rater 1) and 86.6% between direct observations and audio-recordings (rater 1). Inter-rater reliability of audio-recordings (rater 1 vs rater 2) was 81.3%. Further detail is provided in Supplementary File 4.

Fidelity of content:

Fidelity was found to be high in all assessment methods, with a mean score of 81.7% (range of 61.1% - 95.8%) for the audio-recordings, 92.7% (85.2% - 96%) for the self-report and 82.7% (72.1% - 100%) for the direct observations. Table 2 details the fidelity results for each method with scores below 80 (cut-off for 'high' fidelity) shaded. Significant differences between physiotherapists' individual fidelity scores were found. Fidelity scores were also found to differ significantly according to the session category (e.g. the category 'Materials' was delivered with significantly less fidelity than the 'Education' category).

Table 2: Fidelity results calculated from % totals			
	Direct Observations (DO) % (SD)	**Self-Report (SR) % (SD)**	**Audio 1 (AO) % (SD)**
Total % mean fidelity score (SD)	82.7% (10)	92.7% (6.4)	81.7% (12)
% mean fidelity score per session (SD)* **			
1	88.8% (5.24)	95% (4.5)	91.6% (4.5)
2	82.8% (5.7)	92% (6.9)	86.8% (10.5)
3	85.6% (12.9)	96% (4.2)	81.4% (10.6)
4	83.3% (14.4)	90.9% (8.3)	75% (14)
5	74.1% (11.9)	89.4% (8.1)	78.7% (11)
6	82% (9.6)	92.7% (3.78)	74.9% (12)
% mean fidelity score per site (SD)* **			
A (delivered twice, same physiotherapist)	78.7% (7.6)	95% (5.4)	81.3% (11.5)
B (delivered twice, two physiotherapists)	76.7% (5.6)	92.8% (5.3)	71.1% (10)
C (delivered twice, two physiotherapists)	84.8% (11.8)	91.8% (7.7)	84.9% (8.1)
D	87.2% (4)	93.2% (2.9)	87.1% (4.4)
E	83.1% (13)	94.3% (3.8)	88.3% (8.7)

F	72.5%	85.2% (9.6)	72.9% (15)
G	100% (0)	94.7% (4.4)	92.8% (5.4)
% mean fidelity score per category (SD)*			
Materials	72.1% (19.4)	86% (17)	61.1% (29.6)
Introduction and Review	82.6% (16.3)	92.9% (12.8)	76.2% (24.5)
Education	93.3% (8.6)	97.1% (6.6)	95.4% (6.9)
Exercise	80.4% (14)	95.4% (7.1)	82.4% (13)
Review and Planning	77.1% (33)	90.8% (21.6)	69.8% (39.6)
% mean fidelity score per physio (SD)**			
A1	78.8% (7.6)	95.1% (5.4)	81.3% (11.5)
B1	76.1% (7.9)	92.2% (7.2)	72.3% (9.3)
B2	77.5% (0.4)	93.4% (3)	72.6% (12.5)
C1	93.4% (2)	85.2% (4.6)	91% (3.5)
C2	76.2% (11)	98.5% (1.9)	78.8% (6.6)
D1	87.2% (4)	93.2% (2.9)	87.1% (4.4)
E1	83.1% (13)	94.3% (3.8)	88.3% (8.7)
F1	72.5%	85.2% (9.6)	69.8% (14.7)
G1	100% (0)	94.7% (4.4)	95.8% (5.4)

*Significant differences between Categories according to DO (p=0.007), SR (p<0.001) and AO (p<0.001)

** Significant differences between Physiotherapists according to DO (p=0.019), SR (p=0.004) and AO (p<0.001)

***Significant differences between Sites (p<0.001) and between Sessions (p=0.007) according to AO, not significantly different

according to DO and SR.

Shaded areas = scores < 80%

Fidelity of duration:

All methods found a statistically significant difference between the actual duration of the exercise component and its intended duration of 45 minutes (Supplementary File 5). When this was analysed for each individual session for all methods combined, a statistically significant difference ($p<0.001$) was only found for Session 1 between the actual and intended duration. The difference between the actual and intended duration for all other sessions was not found to be statistically significant ($p>0.05$).

Group size – participants in attendance:

The average numbers of participants present for all groups across all sessions was 3, with an SD of 1.3 (Supplementary File 6). Overall, group attendance ranged from one to six participants for any session across both waves. The total number of participants recruited for each group was significantly different between physiotherapists (e.g. seven participants were recruited for the group delivered by physiotherapist F1 compared to only two recruited for the group delivered by D1), as were the numbers of participants present (average group size). The size of groups did not differ significantly between sessions.

Factors associated with fidelity:

Both direct observation and audio-recorded data showed a significant correlation between fidelity scores and the physiotherapists' post-training evaluation scores. Direct observation data also found a significant negative correlation between group sizes and the fidelity scores. Physiotherapist years qualified and experience of delivering groups were found to have significant, negative correlations with fidelity scores for the audio-recorded and self-report data respectively (Table 3).

Table 3: Association between fidelity scores and variables

Variable	Direct Observations	Self-Report	Audio-Recordings	Statistical Test
Group size (number present)	-0.434 (p=0.034)*	-0.215 (p=0.98)	-0.193 (p=0.151)	Spearman's *r* (p-value)
Physiotherapist experience (years qualified)	-0.09 (p=0.676)	-0.186 (p=0.154)	-0.346 (p=0.008)*	

Physiotherapist group experience (years delivering group physiotherapy)	0.171 (p=0.424)	-0.364 (p=0.004)**	0.136 (p=0.312)	
Physiotherapist Post-Training Evaluation Score (%)	0.581 (p=0.003)**	-0.152 (p=0.245)	0.314 (p=0.018)*	
Physiotherapist previous relevant training (yes/no)	*U*=33 (p=0.302)	*U*=201 (p=0.107)	*U* = 243 (p=0.840)	Mann-Whitney U (p-value)
*p is significant at p<0.05, **p<0.005				

Qualitative findings

Inter-rater reliability of coding achieved 81.6% agreement. Overall, physiotherapists felt that they had delivered the programme with good fidelity. All physiotherapists discussed some deviations from the protocol or adaptations made during delivery, i.e. goal-setting was found to be challenging to complete as intended. Other adaptations either concerned difficulties with use of programme materials (e.g. using the projector) as intended or providing additional information during the education content. Five physiotherapists also discussed deviation from protocol in relation to duration, mostly during the first session, with one stating that her *'time management around the education wasn't always exactly what it should have been'* (B1, transcript line 278-286). In terms of factors that influenced fidelity (i.e. reasons discussed for the aforementioned adaptations and deviations), six themes were found that were structured into three levels of factors – physiotherapist, participant and programme-level (Figure 2).

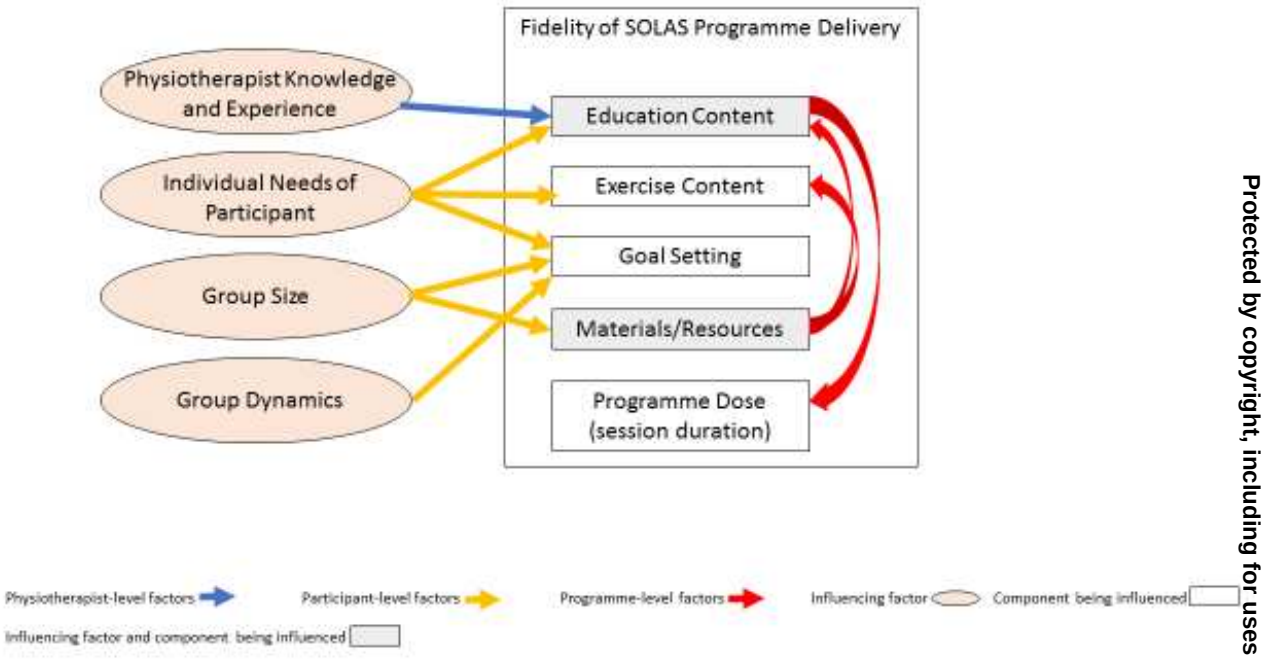

Figure 2: Visual representation of themes of qualitative interviews – factors influencing fidelity of SOLAS delivery

At the physiotherapist level, eight of the nine physiotherapists felt that knowledge of the content of the SOLAS programme facilitated their delivery of the education session and made it easier. Conversely, deviations from protocol within the education content discussed by six physiotherapists were due to the provision of additional information which was influenced by their previous experiences of delivering similar groups (e.g. talking more about pain pathophysiology because of previous classes delivered on this topic). This formed the theme '*Physiotherapist knowledge and experience influenced delivery of SOLAS - education content'.*

At the participant level, five physiotherapists felt that participants' individual needs such as their understanding of content or language literacy levels influenced the delivery of education and exercise components and that adaptation sometimes occurred in response to these needs, creating the theme of '*Individual needs influenced delivery of SOLAS - education, exercise, goal-setting'.* The number of participants present was discussed by seven physiotherapists as another participant-level factor that influenced fidelity of delivery and formed the participant-

level theme of '*Group size influenced delivery of SOLAS - goal-setting, use of materials'.* A further participant-level theme was *'Group dynamics influenced delivery of SOLAS - goal-setting'* as four physiotherapists felt that groups with good dynamics and interaction between participants led to better group discussions and better facilitation of goal setting.

The '*Amount of education content influenced delivery of SOLAS – duration* was a programme-level factor discussed by six physiotherapists, who felt that the amount of education content that was involved in the first session led to more time spent on the education aspect than intended as per protocol. Finally, all nine physiotherapists believed that the good resources (e.g. booklets and handouts, venue space) enhanced and facilitated the delivery of the programme as intended and that occasionally poor or problematic resources (e.g. lack of venue security) negatively influenced the delivery of the programme as intended. This created the theme '*Resources/materials influenced delivery of SOLAS – education and exercise content'.* Exemplary quotes are provided in Table 4.

Table 4: Qualitative findings of factors influencing fidelity results		
Factor level	**Theme**	**Exemplary quote (Physiotherapist code, transcript line)**
Physiotherapist	Physiotherapist knowledge and experience influenced delivery of SOLAS - education content	*'In my previous experience I would have done a lot more actually on the pain side of things…so in my previous class I would have had, you know, maybe one full class on maybe pain perception and, kind of, the influence of emotion and feelings…so I think I would have probably maybe talked a lot more around that pain section than maybe somebody else would have'.* (C1, 75-99)
Participant	Individual needs influenced delivery of SOLAS - education, exercise, goal-setting	'P*eople don't like writing them* [action plans] *there and then you know with pencils given and whatever - yes it's very hard to get people to write down things like that....Where I work there is a lot of people health literacy is very low…so therefore that's a challenge for them…so I tend to be very careful about pushing it out really'* (F1, 141-187)
Participant	Group size influenced delivery of SOLAS – goal-setting, use of materials	*'The only thing I might find a little bit hard would be the goal setting. I suppose you'd - that would be a bit more challenging because you'd have more numbers in the group*' (G1, 118-132)
Participant	Group dynamics influenced delivery of SOLAS - goal-setting	*'People were willing to engage you know as a group, in their goals.....so that made it very easy that we didn't actually have any clients that weren't willing to talk in the group, so it was very much an interactive group*' (E1, 225-231)
Programme	Amount of education content influenced delivery of SOLAS – duration	*'I found the content in week one was nearly too much... by the time I finished talking and ran through the exercises, the hour and a half was finished. And so nobody actually practiced any of the exercises on the first day'* (B2, 96-106)
Programme	Resources/materials	*'The slides didn't work for me this time....You can't lock that room.... once or twice I didn't*

	influenced delivery of SOLAS – education and exercise content	*bring the laptop at all and I just had to print it out, all of the slides on A4 laminate and so we talked all the slides....'* (F1, 207-240) *'I think I only left out maybe three* [exercise] *stations, something like that. Because we didn't have a bouncer and...we didn't have a bed'* (C1, 113-121) *'Nothing but positive feedback for all the content and the-the resources...I just think they complimented the - the education fantastically, I just thought they added much more to the programme than not having these resources.'* (E1, 414-415)

Integrating qualitative and quantitative findings

Each theme was further analysed according to the quantitative data comparing physiotherapists who scored 'high' (i.e. ≥80%; physiotherapists D1, C1, E1 and G1) to those who scored 'moderate' (i.e. ≥50-79%; B1, B2, C2 and F1). Physiotherapist A1 was not included in this analysis as her score was categorised as 'moderate' by direct observations and 'high' by audio-recorded data at 78.7% and 81.3% respectively. A difference between these physiotherapist groups was found in only one theme, '*Group size influenced delivery of SOLAS*'. Physiotherapists who scored higher on the fidelity assessments (average group size of 2.5 participants), believed it was easier to deliver goal-setting as intended with smaller groups. Conversely, physiotherapists with moderate fidelity scores (average group size of 3.7 participants) felt it was harder to facilitate goal-setting as intended with less numbers and believed it would be easier with bigger groups due to better engagement and group discussion. Further details of the triangulation are provided in Table 5 where a meta-matrix was used to compare between findings from both datasets. For the most part, convergence was found between the qualitative and quantitative data, though four qualitative themes relating to influencing factors had no corresponding quantitative data (silence). No areas of discrepancy were found.

Table 5: Triangulation of quantitative and qualitative results – convergence and discrepancy of findings

Outcome of interest	Quantitative fidelity findings		Qualitative interview findings	Convergence/ Discrepancy/ Silence
Fidelity findings:	**Overall fidelity of content scores**	High fidelity (>80%) • Direct Observations: 82.7% • Self-Report: 92.7% • Audio-Recordings: 81.7%	Overall physiotherapists felt that their fidelity was good. Some adaptations and deviations were discussed as occurring in the delivery of the following aspects of the programme: • Goal-setting (Introduction and Review, Review and Planning categories) • Education content • Exercise content • Use of programme materials • Duration of session 1	Convergence
	SOLAS categories scoring below 80% fidelity			
	Materials	Moderate fidelity (50-79%) • Direct Observations: 72.1% • Audio-Recordings: 61.1%		
	Introduction and Review	• Audio-recordings: 76.2%		
	Review and Planning	• Direct Observations: 77.1% • Audio-Recordings: 69.8%		
	Fidelity of duration – sessions significantly different to intended duration			
	Session 1	• Education duration: 58.9'* • Exercise duration: 31.4'*		
Factors influencing	**Correlation between quantitative variables and fidelity scores**		*Theme 1: Physiotherapist knowledge and experience influenced delivery of SOLAS - education content*	Convergence

fidelity findings:	Physiotherapist Experience (years qualified)	• Audio-recordings: -0.333 (p=0.011)***		
	Physiotherapist Group Experience (years delivering group physiotherapy)	• Self-report: -0.430 (p=0.018)***		
	Physiotherapist Post-Training Evaluation Score (%)	• Direct observations: 0.581 (p=0.003)** • Audio-recordings: 0.314 (p=0.018)***	*Theme 1: Physiotherapist knowledge and experience influenced delivery of SOLAS - education content*	Convergence
	Group size (average numbers of participants present)	• Direct observations: -0.434 (p=0.034)***	*Theme 3: Group size influenced delivery of SOLAS – goal-setting, use of materials*	Convergence
	No corresponding quantitative data		*Theme 2: Individual needs influenced delivery of SOLAS - education, exercise, goal-setting*	Silence
			Theme 4: Group dynamics influenced delivery of SOLAS - goal-setting	
			Theme 5: Amount of education content influenced delivery of SOLAS – duration	
			Theme 6: Resources/materials influenced delivery of	

		SOLAS – education and exercise content	
***p<0.001, **p<0.005, ***p<0.05**			

DISCUSSION

The aim of this mixed methods study was to explore and evaluate fidelity of delivery within a feasibility trial of a complex behaviour change intervention using multiple assessment methods. The study found good agreement between researcher-delivered direct observation and audio-recorded fidelity assessment methods, with lower agreement found between provider self-report and researcher-delivered methods. The intervention content was delivered overall with high fidelity, with some variation between physiotherapists and between certain intervention categories. The intervention duration was found to have deviated significantly from intended during the first session only. Subsequently, qualitative interviews with physiotherapists confirmed these fidelity findings. Finally, both qualitative and quantitative data showed that physiotherapists' knowledge and previous experience, as well as the group size, were factors that influenced their fidelity of intervention delivery. The qualitative data contributed further, and postulated additional participant and programme-level factors as aspects that also influenced the overall fidelity results.

Agreement between direct observations and audio-recordings for assessing the fidelity of delivery was found to be excellent. [39] Agreement between both of these methods and provider self-report assessment was lower, as providers consistently rated themselves higher than the independent raters. These findings are perhaps unsurprising, as both direct observations and audio-recordings were rated by the same researcher, and numerous previous studies have shown that providers' subjective assessments of fidelity are often rated higher than independent assessments. [31,32,40] Taking direct observations as the commonly-cited 'gold standard', [11,18,20] these findings reinforce that self-report methods may not be the most accurate method for assessing fidelity in a complex behaviour change study. However, they may still have their place for recording data and also for enhancing fidelity to the protocol by serving as an aide memoire for providers. [41] Although direct observations and audio-recordings have their own limitations [24,42], previous piloting of these assessment methods found that they were feasible and acceptable to physiotherapists. [27] Additionally, the good agreement between audio-recordings and direct observations found in this study suggests that audio-recordings may be a viable alternative with limited resources, as has been done in similar interventions. [43] However, where resources allow, a combination of multiple quantitative methods may provide the most in-depth assessment of fidelity.

One of the key study findings of this study was reinforcing the value of using mixed methods research for the assessment of fidelity. This approach was emphasised in the recent MRC guidelines for conducting process evaluations of complex interventions [6], and is becoming increasingly used in the widespread implementation of evidence-based interventions [44] but does not yet appear to be common practice within fidelity assessments of behaviour change interventions. [10,21] The integration of quantitative and qualitative results enabled the triangulation of findings to provide a better overall picture of the fidelity of the SOLAS intervention and its influencing factors. The importance of the qualitative contribution to answering the 'why' question is evident in the fact that the physiotherapist interviews unearthed strong participant and programme-level factors associated with fidelity results that were not apparent from the quantitative data alone. Whilst this may be predominantly due to the focus of the quantitative analysis on physiotherapist-level variables which were chosen based on existing literature, the participant and programme-level factors identified by this analysis such as group dynamics or amount of programme content may have been difficult to quantitatively analyse to demonstrate association with fidelity results.

This study found that the factors that may influence the fidelity of an interventions' delivery can occur on three levels – provider, participant and programme. Where previous studies have explored factors that have influenced fidelity of intervention delivery, many have focused solely on provider-level factors, demonstrating associations between fidelity and factors such as provider training or skills. [13,19,21,45] The findings of this study have valuable implications for future studies that aim to assess and enhance fidelity of similar interventions as they indicate that planning for fidelity should include considering potential influencing factors at each of these three levels. These results are consistent with recent conclusions by Masterson-Algar et al. in a stroke rehabilitation setting [14] who found that investigating fidelity within clinical trials should also take the individual needs of patients into account, and also concur with the findings of an education-based intervention that found the most common reason for adaptation within intervention delivery was insufficient time. [46]

On the physiotherapist-level, better knowledge of the intervention content and structure was found to positively correlate with quantitative fidelity scores, with a causative link established via the qualitative investigation. This echoes previous findings by Huijg et al. who showed that physiotherapist skill level was one of the most important predictors of fidelity. [21] A more targeted approach to enhancing fidelity in future interventions may

therefore be warranted, such as identifying physiotherapists at higher risk of lower fidelity using post-training evaluation scores and employing more focused fidelity assessment of delivery or further training for these providers, [13] as has been previously employed in similar interventions. [10,43] The results of the study also showed that physiotherapists with more experience of certain aspects tended to emphasise these at the expense of delivering other components as comprehensively as intended in the protocol. These experience-based adaptations invoke the well-established issue of adaptation versus fidelity. For years, research has debated the concept of fidelity versus adaptation, with the case made for both strict fidelity and for modifying interventions. [47] A third view is that both fidelity and adaptation are essential, and achieving an appropriate balance between both can allow an intervention to maximise its effectiveness, while being generalizable and flexible enough to be implementable. [48,49] To achieve this, our fidelity checklists included components that encouraged elements of treatment individualisation (e.g. individualised feedback regarding exercises). However, it may be that these checklists still did not allow for enough individualisation within delivery, an aspect that should be considered by other researchers seeking to conduct similar fidelity assessments.

A limitation of this study was the timing of the interviews, which did not allow a 'pure' convergent/triangulation design. Typically, the qualitative and quantitative methods occur concurrently in this design, [28] however, they were scheduled to take place after physiotherapists had experienced delivery of an entire six-week SOLAS intervention. Although a sequential explanatory design [30] where quantitative results were analysed prior to completing the interviews might have enabled further probing of the factors influencing fidelity, interviews were conducted within two weeks of the intervention completion to minimise recall bias. Due to time constraints it was not possible to have the quantitative data collected and analysed beforehand. Finally, this study mostly focuses on the adherence of delivery (e.g. intervention content and duration) and does not address the *quality* or competence of delivery of SOLAS (e.g. interpersonal or communication style of the physiotherapist), or use of specific BCTs, which is being addressed in a separate publication. This study also does not examine the broader aspects of fidelity such as provider training or participant receipt, as these were beyond the scope of this publication and will be addressed in a future paper.

CONCLUSIONS

In process evaluations and fidelity assessments of large scale complex interventions it is often recommended to complete and report the results of the fidelity assessment before the trial outcomes so as not to bias reporting. [50] Future work will investigate the relationship between this evaluation of fidelity of delivery, the SOLAS feasibility trial outcomes (analysis currently underway) and the evaluation of fidelity of BCT delivery, enabling a potentially more insightful and accurate interpretation of findings. This study also has valuable implications for further research and the overall science of fidelity as it contributes much-needed information to the limited current evidence for the application of fidelity assessment methods within the area of complex behaviour change. The findings have demonstrated how multiple quantitative methods can be used to assess the fidelity of delivery of a complex behaviour change intervention, and that a combination of methods may be most suitable, depending on their acceptability and available resources. We have also shown how the use of a mixed methods approach, integrating both quantitative and qualitative data, provides a more insightful understanding of the factors influencing fidelity.

AUTHORS' CONTRIBUTION

ET conceptualized and designed the study and conducted the data collection and analysis. JM and DAH provided guidance on the design. DAH is the principal investigator of the SOLAS trial and obtained funding for this study within the trial. All authors provided methodological consultation and were involved in substantial contributions to reviewing of the drafts. All authors approved the final version of the manuscript.

ACKNOWLEDGEMENTS

This study was funded as part of Health Research Award HRA_HSR/2012/24 by the Health Research Board of Ireland. The funding source had no other involvement in this publication. The authors would like to thank the physiotherapists and participants for taking part in the study. We would also like to thank Alison Keogh for her help in the inter-rater reliability analysis of the audio-recorded data and physiotherapist interviews and Associate Professor Suzanne Guerin for facilitating the interviews.

CONFLICTS OF INTEREST

None

REFERENCES:

1. Gearing RE, El-Bassel N, Ghesquiere A, Baldwin S, Gillies J, Ngeow E. Major ingredients of fidelity: a review and scientific guide to improving quality of intervention research implementation. *Clin. Psychol. Rev.* 2011;31(1):79-88
2. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implementation Science* 2007;2(40):1-9
3. Borrelli B, Sepinwall D, Ernst D, et al. A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *J. Consult. Clin. Psychol.* 2005;73(5):852-60
4. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337(a1655):1-6
5. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;348
6. Moore GF, Audrey S, Barker M, et al. *Process evaluation of complex interventions: Medical Research Council guidance*, 2015.
7. Hoffmann TC, Walker MF, Langhorne P, Eames S, Thomas E, Glasziou P. What's in a name? The challenge of describing interventions in systematic reviews: analysis of a random sample of reviews of non-pharmacological stroke interventions. *BMJ Open* 2015;5(11)
8. Schinckus L, Broucke S, Housiaux M. Assessment of implementation fidelity in diabetes self-management education programs: a systematic review. *Patient. Educ. Couns.* 2014;96(1):13-21
9. Toomey E, Currie-Murphy L, Matthews J, Hurley DA. Implementation fidelity of physiotherapist-delivered group education and exercise interventions to promote self-management in people with osteoarthritis and chronic low back pain: A rapid review Part II. *Man. Ther.* 2015;20(2):287-94
10. Mars T, Ellard D, Carnes D, Homer K, Underwood M, Taylor SJC. Fidelity in complex behaviour change interventions: a standardised approach to evaluate intervention integrity. *BMJ Open* 2013;3(11)
11. French SD, Green SE, Francis JJ, et al. Evaluation of the fidelity of an interactive face-to-face educational intervention to improve general practitioner management of back pain. *BMJ Open* 2015;5(7)
12. Perepletchikova F, Hilt LM, Chereji E, Kazdin AE. Barriers to implementing treatment integrity procedures: survey of treatment outcome researchers. *J Consult Clin Psychol* 2009;77(2):212-8
13. Wang B, Stanton B, Deveaux L, et al. Factors influencing implementation dose and fidelity thereof and related student outcomes of an evidence-based national HIV prevention program. *Implement Sci* 2015;10(1):44
14. Masterson-Algar P, Burton CR, Rycroft-Malone J, Sackley CM, Walker MF. Towards a programme theory for fidelity in the evaluation of complex interventions. *Journal of Evaluation in Clinical Practice* 2014;20(4):445-52
15. Langford R, Bonell C, Jones H, Campbell R. Obesity prevention and the Health promoting Schools framework: essential components and barriers to success. *The International Journal of Behavioral Nutrition and Physical Activity* 2015;12:15
16. Mowbray CT, Holter MC, Teague GB, Bybee D. Fidelity Criteria: Development, Measurement, and Validation. *American Journal of Evaluation* 2003;24(3):315-40
17. Poltawski L, Norris M, Dean S. Intervention fidelity: Developing an experience-based model for rehabilitation research. *J. Rehabil. Med.* 2014;46(7):609-15

18. Bellg AJ, Borrelli B, Resnick B, et al. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol.* 2004;23(5):443-51
19. Lorencatto F, West R, Christopherson C, Michie S. Assessing fidelity of delivery of smoking cessation behavioural support in practice. *Implementation Science* 2013;8(40)
20. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J. Public Health Dent.* 2011;71(s1):S52-S63
21. Huijg JM, Dusseldorp E, Gebhardt WA, et al. Factors associated with physical therapists' implementation of physical activity interventions in The Netherlands. *Phys. Ther.* 2015;95(4):539-57
22. McKenna JW, Flower A, Ciullo S. Measuring Fidelity to Improve Intervention Effectiveness. *Intervention in School and Clinic* 2014;50(1):15-21
23. Resnicow K, Davis M, Smith M, et al. How best to measure implementation of school health curricula: a comparison of three measures. *Health Education Research* 1998;13(2):239-50
24. Breitenstein SM, Gross D, Garvey CA, Hill C, Fogg L, Resnick B. Implementation fidelity in community-based interventions. *Res. Nurs. Health* 2010;33(2):164-73
25. Hurley DA, Hall AM, Currie-Murphy L, et al. Theory-driven group-based complex intervention to support self-management of osteoarthritis and low back pain in primary care physiotherapy: protocol for a cluster randomised controlled feasibility trial (SOLAS). *BMJ Open* 2016;6(1):e010728
26. Hurley DA, Murphy LC, Hayes D, et al. Using intervention mapping to develop a theory-driven, group-based complex intervention to support self-management of osteoarthritis and low back pain (SOLAS). *Implementation Science* 2016;11(1):1-29
27. Toomey E, Matthews J, Guerin S, Hurley DA. Development of a Feasible Implementation Fidelity Protocol Within a Complex Physiotherapy-Led Self-Management Intervention. *Phys. Ther.* 2016
28. Creswell JW, Plano Clark VL. *Designing and Conducting Mixed Methods Research*. 2 ed. Thousand Oaks, CA: SAGE Publications, 2011.
29. Barbour RS. The case for combining qualitative and quantitative approaches in health services research. *J. Health Serv. Res. Policy* 1999;4(1):39-43
30. Ivankova NV, Creswell JW, Stick SL. Using Mixed-Methods Sequential Explanatory Design: From Theory to Practice. *Field Methods* 2006;18(1):3-20
31. Hardeman W, Michie S, Fanshawe T, Prevost AT, McLoughlin K, Kinmonth AL. Fidelity of delivery of a physical activity intervention: Predictors and consequences. *Psychol. Health* 2008;23(1):11-24
32. Breitenstein SM, Fogg L, Garvey C, Hill C, Resnick B, Gross D. Measuring Implementation Fidelity in a Community-Based Parenting Intervention. *Nursing research* 2010;59(3):158-65
33. Perepletchikova F, Kazdin AE. Treatment Integrity and Therapeutic Change: Issues and Research Recommendations. *Clinical Psychology: Science and Practice* 2005;12(4):365-83
34. Garbacz L, Brown D, Spee G, Polo A, Budd K. Establishing Treatment Fidelity in Evidence-Based Parent Training Programs for Externalizing Disorders in Children and Adolescents. *Clin Child Fam Psychol Rev* 2014;17(3):230-47
35. Braun V, Clarke V. Using thematic analysis in psychology. *Qual. Res. Psychol.* 2006;3(2):77-101
36. Guerin S, Hennessy E. Pupils' definitions of bullying. *Eur J Psychol Educ* 2002;17(3):249-61
37. Wendler MC. Triangulation using a meta-matrix. *J. Adv. Nurs.* 2001;35(4):521-5
38. Farmer T, Robinson K, Elliott SJ, Eyles J. Developing and implementing a triangulation protocol for qualitative health research. *Qualitative Health Research* 2006;16(3):377-94
39. Fleiss J. Reliability of Measurement. In: Fleiss J, ed. The design and analysis of clinical experiments. New York: Wiley, 1986:1-32.
40. Carroll K, Nich C, Rounsaville B. Utility of Therapist Session Checklists to Monitor Delivery of Coping Skills Treatment for Cocaine Abusers. *Psychotherapy Research* 1998;8(3):307-20

41. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Publ Health Dentistry* 2011;71:S52 - S63
42. Leventhal H, Friedman MA. Does establishing fidelity of treatment help in understanding treatment efficacy? Comment on Bellg et al. (2004). *Health Psychol.* 2004;23(5):452-6
43. Pincus T, Anwar S, McCracken LM, et al. Delivering an Optimised Behavioural Intervention (OBI) to people with low back pain with high psychological risk; results and lessons learnt from a feasibility randomised controlled trial of Contextual Cognitive Behavioural Therapy (CCBT) vs. Physiotherapy. *BMC Musculoskelet. Disord.* 2015;16:147
44. Palinkas LA, Aarons GA, Horwitz S, Chamberlain P, Hurlburt M, Landsverk J. Mixed Method Designs in Implementation Research. *Administration and Policy in Mental Health* 2011;38(1):44-53
45. Taylor CA, Shaw RL, Dale J, French DP. Enhancing delivery of health behaviour change interventions in primary care: a meta-synthesis of views and experiences of primary care nurses. *Patient Educ. Couns.* 2011;85(2):315-22
46. Hill LG, Maucione K, Hood BK. A focused approach to assessing program fidelity. *Prev Sci* 2007;8(1):25-34
47. Dane AV, Schneider BH. Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin. Psychol. Rev.* 1998;18(1):23-45
48. Castro FG, Barrera M, Jr., Martinez CR, Jr. The cultural adaptation of prevention interventions: resolving tensions between fidelity and fit. *Prev Sci* 2004;5(1):41-5
49. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology* 2008;41(3-4):327-50
50. Oakley A, Strange V, Bonell C, Allen E, Stephenson J. Health services research: process evaluation in randomised controlled trials of complex interventions. *BMJ* 2006;332:413

Supplementary Files

Supplementary File 1: Quantitative fidelity checklists used

INTERVENTION SESSION CHECKLIST (OBSERVATION/AUDIO-RECORDING):

Cover sheet (completed for each session)

Date:	
Venue:	
Physiotherapist Name:	
Other staff involved:	Name(s): Role: (e.g. set-up/delivery/support)
Session number (tick):	1☐ 2☐ 3☐ 4☐ 5☐ 6☐
Start time (class):	
Finish time (class):	
Time spent on education (mins):	
Time spent on exercise (mins):	
Method of Observation:	In-vivo/Audio-recording
Adverse event(s)/issue(s) (circle): **Y/N**	If yes give brief details:

Deviations from protocol/proscribed components delivered? **General notes on fidelity of session:**	Yes/No Details:				
INTERVENTION COMPONENT CHECKLIST:		**YES (2)**	**NO (0)**	**ATTEMPTED (1)**	**N/A**
Session 1:					
Materials					
	Intervention folder given to participants	☐	☐	☐	
	Use of **participant intervention folder** actively facilitated **throughout** session	☐	☐	☐	
	Name stickers/badges given to participants	☐	☐	☐	
	Pens offered/provided	☐	☐	☐	
	Powerpoint slides used	☐	☐	☐	
Introduction					
	Introductions/welcome made	☐	☐	☐	
	Set **clear expectations - aims**, **content** and **structure** of programme addressed	☐	☐	☐	
	Rationale for **weekly attendance** provided	☐	☐	☐	
Education					

Rationale for **self-management** provided	☐	☐	☐	
Posture addressed	☐	☐	☐	
Cycle of change addressed	☐	☐	☐	
Disease characteristics, prevalence and causes of OA/CLBP addressed	☐	☐	☐	
Recommended activity levels/benefits of exercise addressed	☐	☐	☐	
Individual reflection on activity/recommendations facilitated	☐	☐	☐	
Goal setting introduced	☐	☐	☐	
Action planning introduced	☐	☐	☐	
Attention drawn to **Action Plan sheets** within intervention folder, **encouraged use**	☐	☐	☐	
Attention drawn to **Activity Diaries** within intervention folder, **encouraged use**	☐	☐	☐	
Participants given a chance/encouraged to **contribute to discussion**	☐	☐	☐	
Exercise				
Room set up for exercise (equipment, sheets)	☐	☐	☐	
Protocol exercises **demonstrated**	☐	☐	☐	
Attention drawn to **Exercise Programme Diary** within intervention folder, **encouraged use**	☐	☐	☐	
Rationale for **exercises** provided	☐	☐	☐	
Participants given a chance to **attempt and practice protocol** exercises	☐	☐	☐	
Individual feedback provided	☐	☐	☐	☐
Review and Planning				
Session review - **activity levels and goal setting recap**	☐	☐	☐	
Proscribed components delivered?	☐(-2)	☐		
Total score (Yes = 2, Attempted = 1, No =0)				
Overall Adherence score				

Session 2:

INTERVENTION COMPONENT CHECKLIST:	**YES (2)**	**NO (0)**	**ATTEMPTED (1)**	**N/A**

Materials					
Use of **participant intervention folder** actively facilitated **throughout** session	☐	☐	☐		
Name stickers/badges given to participants	☐	☐	☐		☐
Pens offered/provided	☐	☐	☐		
Pedometers offered	☐	☐	☐		
Powerpoint slides used	☐	☐	☐		
Recap and Review					
Welcome back made	☐	☐	☐		
Review of previous week **activities/action plans**	☐	☐	☐		
Education					
Set clear expectations – content of session outlined	☐	☐	☐		
Activity-rest cycle and **pacing** explained	☐	☐	☐		
Individual reflection on **current pacing/activity-rest** facilitated	☐	☐	☐		
Factors influencing pain addressed	☐	☐	☐		
Goal setting facilitated	☐	☐	☐		
Action planning facilitated	☐	☐	☐		
Action Plan sheets use encouraged/facilitated	☐	☐	☐		
Attention drawn to **Walking/Activity Diary** within intervention folder, **encouraged use**	☐	☐	☐		
Participants given a chance/encouraged to **contribute to discussion**	☐	☐	☐		
Exercise					
Room set up for exercise (equipment, sheets)	☐	☐	☐		
Protocol exercises **demonstrated**	☐	☐	☐		☐
Exercise Programme Diary use encouraged	☐	☐	☐		
Rationale for **exercises** provided	☐	☐	☐		
Participants given a chance to **attempt and practice protocol** exercises	☐	☐	☐		
Individual feedback provided	☐	☐	☐		☐
Session Planning and Review					
Session review - **goal setting and action planning recap**	☐	☐	☐		
Proscribed components delivered?	☐(-2)	☐			
Total score (Yes = 2, Attempted = 1, No =0)					

Overall Adherence score				
Session 3:				
INTERVENTION COMPONENT CHECKLIST:	**YES (2)**	**NO (0)**	**ATTEMPTED (1)**	**N/A**
Materials				
Use of **participant intervention folder** actively facilitated **throughout** session	☐	☐	☐	
Tape measures offered	☐	☐	☐	
Pens offered/provided	☐	☐	☐	
Powerpoint slides used	☐	☐	☐	
Recap and Review				
Previous week **Activity Action Plan** reviewed	☐	☐	☐	
Problem-solving of previous week **Activity Action Plan**	☐	☐	☐	☐
Action Plan sheets use encouraged/facilitated	☐	☐	☐	
Education				
Set clear expectations – content of session outlined	☐	☐	☐	
Obesity and effect on pain addressed	☐	☐	☐	
Balance between weight/activity addressed	☐	☐	☐	
Individual reflection on **weight/activity** balance facilitated	☐	☐	☐	
Skills for maintaining healthy weight addressed (e.g. waist measurement, food diary)	☐	☐	☐	
Attention drawn to **Healthy Eating Booklet** within intervention folder, **encouraged use**				
Attention drawn to **Walking/ Activity Diary** within intervention folder, **encouraged use**	☐	☐	☐	
Attention drawn to **Food & Exercise Diary** within intervention folder, **encouraged use**	☐	☐	☐	
Participants given a chance/encouraged to **contribute to discussion**	☐	☐	☐	
Exercise				
Room set up for exercise (equipment, sheets)	☐	☐	☐	
Protocol exercises **demonstrated**	☐	☐	☐	☐

Exercise Programme Diary use encouraged	☐	☐	☐	
Rationale for **exercises** provided	☐	☐	☐	
Participants given a chance to **attempt and practice protocol** exercises	☐	☐	☐	
Individual feedback provided	☐	☐	☐	☐
Session Planning and Review				
Session review - **goal setting and action planning** recap integrating **food and exercise diary**	☐	☐	☐	
Proscribed components delivered?	☐(-2)	☐		
Total score (Yes = 2, Attempted = 1, No =0)				
Overall Adherence score				

Session 4:

INTERVENTION COMPONENT CHECKLIST:	**YES (2)**	**NO (0)**	**ATTEMPTED (1)**	**N/A**
Materials				
Use of **participant intervention folder** actively facilitated **throughout** session	☐	☐	☐	
Pens offered/provided	☐	☐	☐	
Powerpoint slides used	☐	☐	☐	
Midway Recap and Review				
Previous week **Activity Action Plan** reviewed	☐	☐	☐	
Problem-solving of previous week **Activity Action Plan**	☐	☐	☐	☐
Action Plan sheets use encouraged/facilitated	☐	☐	☐	
Walking/ Activity Diary use encouraged	☐	☐	☐	
Education				
Set clear expectations – content of session outlined	☐	☐	☐	
Rationale for using **pain relief** given (e.g. **pain pathway** explained)	☐	☐	☐	
Methods of pain relief addressed (e.g. medication, heat/ice, TENS/acupuncture)	☐	☐	☐	
Individual reflection on use of **pain relief** facilitated	☐	☐	☐	

Participants given a chance/encouraged to **contribute to discussion**	☐	☐	☐	
Exercise				
Room set up for exercise (equipment, sheets)	☐	☐	☐	
Protocol exercises **demonstrated**	☐	☐	☐	☐
Exercise Programme Diary use encouraged	☐	☐	☐	
Rationale for **exercises** provided	☐	☐	☐	
Participants given a chance to **attempt and practice protocol** exercises	☐	☐	☐	
Individual feedback provided	☐	☐	☐	☐
Session Planning and Review				
Session review - **goal setting and action planning** recap	☐	☐	☐	
Proscribed components delivered?	☐(-2)	☐		
Total score (Yes = 2, Attempted = 1, No =0)				
Overall Adherence score				

Session 5:

INTERVENTION COMPONENT CHECKLIST:	**YES (2)**	**NO (0)**	**ATTEMPTED (1)**	**N/A**
Materials				
Use of **participant intervention folder** actively facilitated **throughout** session	☐	☐	☐	
Pens offered/provided	☐	☐	☐	
Relaxation CD offered	☐	☐	☐	
Powerpoint slides used	☐	☐	☐	
Recap and Review				
Previous week **Activity Action Plan** reviewed	☐	☐	☐	
Problem-solving of previous week **Activity Action Plan**	☐	☐	☐	☐
Action Plan sheets use encouraged/facilitated	☐	☐	☐	
Walking/ Activity Diary use encouraged	☐	☐	☐	
Education				
Set clear expectations – content of session outlined	☐	☐	☐	

Recognising and managing flare-ups addressed	☐	☐	☐	
Individual reflection about **flare-ups** facilitated	☐	☐	☐	
Effect of mood on pain addressed	☐	☐	☐	
Participants given a chance/encouraged to **contribute to discussion**	☐	☐	☐	
Exercise				
Room set up for exercise (equipment, sheets)	☐	☐	☐	
Protocol exercises **demonstrated**	☐	☐	☐	☐
Exercise Programme Diary use encouraged	☐	☐	☐	
Rationale for **exercises** provided	☐	☐	☐	
Participants given a chance to **attempt and practice protocol** exercises	☐	☐	☐	
Individual feedback provided	☐	☐	☐	☐
Relaxation Session				
Relaxation techniques explained and **practiced**	☐	☐	☐	
Session Planning and Review				
Session review - **goal setting and action planning** recap with **integration** of **relaxation techniques**	☐	☐	☐	
Proscribed components delivered?	☐(-2)	☐		
Total score (Yes = 2, Attempted = 1, No =0)				
Overall Adherence score				

Session 6:

INTERVENTION COMPONENT CHECKLIST:	**YES (2)**	**NO (0)**	**ATTEMPTED (1)**	**N/A**
Materials				
Use of **participant intervention folder** actively facilitated **throughout** session	☐	☐	☐	
Pens offered/provided	☐	☐	☐	
Handouts on **local resources and supports** provided	☐	☐	☐	
Powerpoint slides used	☐	☐	☐	
Recap and Review				
Previous week **Activity Action Plan** reviewed	☐	☐	☐	
Problem-solving of previous week **Activity Action Plan**	☐	☐	☐	☐

Action Plan sheets use encouraged/facilitated	☐	☐	☐	
Walking/ Activity Diary use encouraged	☐	☐	☐	
Education				
Set clear expectations – content of session outlined	☐	☐	☐	
Core skills of programme reviewed	☐	☐	☐	
Aims of long-term self-management addressed	☐	☐	☐	
Local resources and supports discussed	☐	☐	☐	
Participants given a chance/encouraged to **contribute to discussion**	☐	☐	☐	
Exercise				
Room set up for exercise (equipment, sheets)	☐	☐	☐	
Protocol exercises **demonstrated**	☐	☐	☐	☐
Exercise Programme Diary use encouraged	☐	☐	☐	
Rationale for **exercises** provided	☐	☐	☐	
Participants given a chance to **attempt and practice protocol** exercises	☐	☐	☐	
Individual feedback provided	☐	☐	☐	☐
Session Planning and Review				
Session review – **long term goal setting and action planning** recap	☐	☐	☐	
Activity diaries **use recorded** (if willing)	☐	☐	☐	
Proscribed components delivered?	☐(-2)	☐		
Total score (Yes = 2, Attempted = 1, No =0)				
Overall Adherence score				

Yes = Fully addressed by the Physiotherapist – could not do more

Attempted = Reasonable attempt made to address this – could do more

No = No attempt made to address this

INTERVENTION SESSION CHECKLIST (SELF-REPORT)

Self-report checklists covered similar components to observation and audio as provided above - different cover sheet (completed for each session) provided below

PCCC Site:	Date:	Class: circle 1 2 3 4 5 6	Physiotherapist Name:

PERSONNEL

Other staff involved in setting up class:	Yes ☐ No ☐	Names/Staff Grade:
Other staff involved in providing class:	Yes ☐ No ☐	Names/Staff Grade:

ATTENDANCE

	Present	Absent
Number of Clients:		
Names of non-attenders:	Reasons for non-attendance [if known]	
1. 2. 3. 4.		

CLASS PREPARATION

Time to review materials [mins]:	Time to set up class [mins]:	Time to take down class [mins]:
Start Time:	End Time:	Comments:

CLASS DELIVERY

	Education	Exercise
Time to deliver [mins]:		
Comments:		
Deviations from protocol: Content/time	Yes ☐ No ☐	Yes ☐ No ☐
If 'yes' give details and reason[s]		

TREATMENT-RELATED EVENT – any unforeseen event/issue should be reported to Deirdre Hurley

During class: Yes ☐ No ☐	**After class:** Yes ☐ No ☐	**Reported to Deirdre** Yes ☐ No ☐

When patient is discharged please give this completed form to UCD Research Physiotherapist or scan and email to physiostudy@ucd.ie

Supplementary File 2: Physiotherapist baseline characteristics

	Wave 1				Wave 2					
Physiotherapy Site	A	B	C	D	A	B	C	E	F	G
Physiotherapist	A1	B1	C1	D1	A1	B2	C2	E1	F1	G1
Gender	F	F	F	M	F	F	F	F	F	F
Experience (years qualified)	5	25	8	12	5	25	10	6	19	11
Group experience (years delivering group physiotherapy)	1	3	8	2	1	10	3	1	15	7
Post-training evaluation score (%)	74%	81.5%	90.5%	83.4%	74%	88.5%	81.8%	88.2%	78.1%	94.4%
Previous training in similar interventions (Y/N)	N	Y	Y	Y	N	Y	Y	Y	Y	Y
F= female, M=male, Y=yes, N=no										

Supplementary File 3: Physiotherapist interviews topic guide

SOLAS Programme Semi-Structured Interviews with Physiotherapists

Interview Schedule

Ice-breaker – describe your experience in running group-based programmes for these or other populations before this study. How many weeks of the SOLAS programme did you deliver?

Therapist Views on Experience of Delivering the Programme

- What are your overall impressions of the programme having delivered it for first/second time?
    - Content overall and week by week –education and exercise components – time for each- managing group dynamics – mixed ability/diagnosis and ages of clients
- What aspects did you find easy/challenging to deliver? [content/communication/behaviour change/goal setting/action planning – these were emphasised in training]
- Views on feasibility of class size of up to 8 clients with one physiotherapist to deliver? [none delivered a class of this size – their views on running small numbers v the target of 8]
- How well do you feel you delivered the programme as intended from the training received?
    - Following the slide content/script – was it difficult/did you want to edit - which parts?
    - Content/needs supportive delivery style/ behaviour change techniques
    - Views on giving advice/setting goals with patients and following through and being needs supportive (using SDT) e.g. using non-controlling language, enabling patient input and choice; providing positive and personalised feedback to patients?
    - Difficult/constrained by research?
- For the aspects not delivered as planned from training give reasons –
    - prompt on potential barriers to delivery [the availiability of resources (e.g. staffing, suitable venue, administrative staff, time constraints); appropriate patient selection and screening, patient uptake and engagement with programme and the potential need to individualise treatment within group]
- How much additional work did delivering the programme and participating in this study place upon you? [Specify – preparation time – reading the manual and supplementary materials, setting up the venue, time to deliver – on top of other work, completion of treatment record forms after each class, completion of post training questionnaires]?
    - Is this acceptable? What modifications would you suggest for future waves?
- What are your impressions of the resources provided to you to support delivery of the programme? [Training Manual /Intervention Manual/Intervention Slides/SOLAS poster]
    - How much have you used them? What was useful/not useful in terms of helping you deliver the programme as intended?
    - Views on continuing to use powerpoint versus flip chart or handouts only?
    - Suggestions for modifications for future waves
- What are your impressions of the venue in terms of its suitability for delivering the programme? [Prompt – accessibility, space, equipment for delivering education and exercise component number of stations, sound quality to allow communication to the

group/individuals during exercises

Views on sustainability of the Programme

- What is your opinion of the feasibility of delivering the programme in the future within the study/within normal practice?
- Would you deliver the programme outside the study?
- How would you deliver it? What materials would you use? What would you leave out? Who would you deliver it to?

Views on Research Elements of the Programme

- Views on research elements of the programme [pre and post training questionnaires, direct observation during classes, audio recordings, treatment record forms]
    - Intrusiveness/time/feasibility Any suggestions for modifications for future waves
- Views on level and modes of communication with the research team throughout the study from training to completion of this wave
    - Suggestions for modifications for future waves

Supplementary File 4: Agreement between assessment methods				
	Direct Observation v Self-Report (n=24)	Direct Observation v Audio-Recordings (n=24)	Audio-Recordings v Self-Report (n=60)	Audio-Recordings Rater 1 v Rater 2 (Inter-rater reliability) (n=12)
Overall agreement	74.6%	86.6%	75.4%	84.6%
% Agreement per category:				
Materials	74.5%	82.67%	70.1%	84.6%
Introduction and Review	65.3%	86.5%	57%	81.6%
Education	84.6%	90.3%	87.3%	76.7%
Exercise	70.8%	86%	78%	83.3%
Review and Planning	50%	76.2%	46%	100%