# Development of the Brief Addiction Therapist Scale (BATS): a tool for evaluating therapist delivery of psychological therapies used in routine practice for alcohol and drug use problems


Tober, G., Crosby, H.F., Latchford, G., and Bewick, B.M.

October 2018

## AUTHOR DETAILS

Gillian Tober[1,2]
Email: gilliantober@yahoo.co.uk
Tel: +44 (0)113 855 9591

Helen F. Crosby[1,2]
Email: H.F.Crosby@leeds.ac.uk, bats@leeds.ac.uk
Tel: +44 (0)113 343 1898

Gary Latchford[2]
Email: G.Latchford@leeds.ac.uk
Tel: +44 (0)113 343 2736

Bridgette M. Bewick[2]
Email: B.M.Bewick@leeds.ac.uk
Tel: +44 (0)113 343 0809

[1]Leeds Addiction Unit Training and Research, The Coach House, 19 Springfield Mount, Leeds, UK
[2]Leeds Institute of Health Sciences, Level 10 Worsley Building, School of Medicine, University of Leeds, Clarendon Way, Leeds, UK

## ACKNOWLEDGEMENTS

We acknowledge the support from the University of Leeds, and Leeds and York Partnership NHS Foundation Trust (LYPFT). We thank Magpie for their contribution in designing the BATS logo. We are grateful to the participants who originally gave consent for their sessions to be used for research, and for the participants in the current study. The BATS was funded by Alcohol Research UK, the Society for the Study of Addiction (SSA), and the University of Leeds Max Hamilton Psychiatric Research Fund. The study formed part of HC's PhD.

This report was funded by Alcohol Research UK. Alcohol Research UK and Alcohol Concern merged in April 2017 to form a major independent national charity, working to reduce the harms caused by alcohol.

Read more reports at: www.alcoholresearchuk.org

*Opinions and recommendations expressed in this report are those of the authors.*

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ADAPTA | Addressing Drinking Among Patients: comparing Two Approaches<br>(An alcohol-focused intervention versus a healthy living intervention for problem drinkers identified in a general hospital setting) |
| AESOPS | Alcohol: Evaluating Stepped care in Older Populations Study |
| AF | Alcohol Focused intervention |
| BA | Brief Advice |
| BATS | Brief Addiction Therapist Scale |
| BCC | Behaviour Change Counselling |
| CBT | Cognitive Behavioural Therapy |
| DAS | Drug and Alcohol Service |
| HL | Health Living focused intervention |
| ICC | Intraclass correlation coefficient |
| MET | Motivational Enhancement Therapy |
| MISTS | Motivational Interviewing Supervision Training Scale |
| MTRS | MATCH Tape Rating Scale |
| PRS | Process Rating Scale |
| RP | Routine Practice |
| SAS | (NHS) Specialist Addictions Service |
| SBNT | Social Behaviour and Network Therapy |
| UKATT | United Kingdom (UK) Alcohol Treatment Trial |
| YACS | Yale Adherence Competence Scale |

# CONTENTS

EXECUTIVE SUMMARY ..... 1

BACKGROUND ..... 3

FINDINGS ..... 9

DISCUSSION ..... 16

REFERENCES ..... 17

APPENDIX 1: BATS version 1 ..... 20

# EXECUTIVE SUMMARY

## Background

Research on treatment for alcohol and drug misuse suggests there are key active mechanisms common to a number of approaches and protocols. Increasingly, treatment effectiveness studies have sought to identify how these 'active ingredients' of treatment are used, and contribute to change, in treatment settings. However, adherence to treatment protocols, and practice consistent with these mechanisms (sometimes termed 'treatment integrity' or 'fidelity'), remains a challenge. This has implications for both practitioners and those responsible for their supervision. The assessment and monitoring of treatment fidelity requires reliable and valid measures which can be used in different settings.

## Aim

The study aimed to develop the Brief Addiction Therapist Scale (BATS): a tool for evaluating the delivery of psychological therapies for alcohol and drug use problems in routine practice.

## Method

A literature review was conducted to identify fidelity measures that evaluated the delivery of psychological therapies for alcohol and drug use problems. This found 26 relevant measures comprising 783 items. The items were coded using thematic analysis, and 18 exemplar items were identified.

Following the literature review, a modified three-round Delphi survey was used to obtain a consensus of opinion among selected experts on the content of the scale. 12 participants took part in rounds one and two, and 10 took part in round three. At the conclusion of round three, group agreement on the scale items and definitions was reached. This formed the content of the first version of the scale. Consultation with therapists, who are the intended users of the scale, improved the guidance notes. Recording were made of therapy sessions in routine addiction services, and these were rated using the new scale. These ratings compared with ratings from other measures to ascertain concurrent validity. Independent double rating showed good to excellent inter-rater reliability.

The BATS was then used to rate therapy sessions, which had been recorded either as part of routine practice by therapists working at an addiction service, or as part of three randomised controlled trials: UKATT (UKATT Research Team, 2005), AESOPS (Watson et al., 2013a), and ADAPTA (Watson et al., 2013b).

## Conclusion

The project led to the development of a brief, evidence-based tool for monitoring and evaluating the delivery of psychological therapies in routine practice. The BATS is trans-theoretical, and applicable to the range of widely used therapies in addiction. The BATS is being used to support supervision at an NHS addiction service demonstrating its utility in routine practice and its potential to have an impact on service delivery.

## BACKGROUND

Studies on the effectiveness of treatments for alcohol and drug misuse have shown the equivalence of a number of well-structured cognitive, behavioural and network based treatments. Consequently, there has been growing interest in understanding the common ingredients that form the active mechanisms of therapeutically driven behaviour change. Protocols for such effective therapies include: motivational enhancement therapy (MET), social behaviour and network therapy (SBNT) and cognitive-behavioural therapy (CBT) (e.g. Miller et al., 1992; Copello et al., 2002; Carroll, 1998). The putative active components have been correlated with outcome data in order to identify those that are most effective. The preferred content and style of delivery is specified in treatment manuals, which provide the basis for monitoring and assessing fidelity (Schoenwald et al., 2011).

Treatment fidelity, sometimes termed treatment integrity, has three components:

- Treatment adherence
- Therapist competence
- Treatment differentiation (Schoenwald and Garland, 2013).

*Adherence* refers to the extent to which therapies are delivered as described by treatment protocols and manuals. *Competence* refers to the level of skill shown by the therapist in delivering the therapies (e.g. Waltz et al., 1993). Most definitions of treatment fidelity emphasise these first two components. The third aspect, treatment *differentiation*, refers to the question of whether the therapies are distinguishable from one another in clinical trials. A breakdown in any one of these components may compromise treatment fidelity and threaten the internal validity of studies (Perepletchikova and Kazdin, 2005).

The assessment of treatment fidelity requires reliable and valid measures. Treatments shown to be effective in research trials need to be delivered with fidelity in routine practice. Fidelity monitoring supports the delivery of evidence-based practice and promotes continuous quality improvement (Aarons et al., 2011). A number of scales for measuring fidelity have been developed:

The *MATCH Tape Rating Scale* (MTRS) was one of the earliest scales to be developed (Carroll et al., 1998). Adapted from assessment methods used in the National Institute of Mental Health Treatment of Depression Collaborative Research Program (Hill et al., 1992), the scale incorporated Likert-type items to assess adherence, and to differentiate between, the treatments that had been previously compared in Project MATCH (Project MATCH Research Group, 1997). The MTRS distinguished between trial treatments but did not measure competence.

The *Yale Adherence and Competence Scale* (YACS) is a more comprehensive measure designed to assess adherence and competence in the delivery of multiple therapies commonly used for substance use disorders (Carroll et al., 2000). As the YACS is a research tool, its use in other settings (for example, training and supervision contexts) is limited (Madson and Campbell, 2006).

The *UKATT Process Rating Scale* was developed and validated in the UK context to measure the delivery and distinctiveness of Motivational Enhancement Therapy and Social Behaviour and Network Therapy in the UK Alcohol Treatment Trial (Tober et el. 2008).

The *Motivational Interviewing Supervision Training Scale* (MISTS) was developed to monitor the training and supervision of therapists practising motivational interviewing (Madson et al., 2005). This has yet to be evaluated outside of the research environment (Madson and Campbell, 2006).

Measures of fidelity that have been developed in the context of research trials are generally unsuitable for use in routine practice because they are either too long (e.g. Carroll et al., 2000, Tober et al. 2008) or are specific to one therapy (e.g. Madson et al., 2005). However, they can be adapted for use in routine practice – particularly to inform clinical supervision. Therapists tend to be eclectic in their choice of treatment methods and theoretical orientation (Raistrick et al., 2006) and tailor interventions to meet individual client needs (Norcross and Wampold, 2011). To have utility in routine practice, a trans-theoretical scale is therefore needed which can evaluate the delivery of a range therapies across different settings.

# METHOD

## Generating an item pool

A systematic search of the literature was conducted to identify fidelity measures that evaluated the delivery of psychological therapies for alcohol and drug use problems. The literature was searched for articles describing the development and/or validation of relevant fidelity measures. Searches were conducted in Medline, Embase, and PsycINFO. Reference lists of articles included in the review were also searched. Citations from the database and reference list searches were downloaded into an endnote library (EndNote X7). HC screened the article titles and abstracts. Full manuscripts of potentially relevant articles were retrieved and assessed for inclusion by the research team (HC, GL, BB, and GT) using criteria listed in Table 1. Finally, measures that evaluate the delivery of psychological therapies for alcohol and drug use problems were extracted from those articles which met inclusion criteria.

Items from the identified fidelity measures were coded thematically. The analysis was based on guidelines provided by Braun and Clarke (2006). There were 6 phases (see Table 2). Exemplar items were developed from each theme, and were used as a basis for generating an initial pool of items to be used in round one of the Delphi survey.

## Refining the scale content

A modified three-round Delphi survey was used to obtain a consensus of opinion among selected experts on the content of the BATS (items, item scoring and item definitions). Nineteen participants were selected for their expertise in the areas of addiction and psychotherapy. Data were collected using a series of questionnaires,

or rounds, in which participant responses from one round were used to inform the next round. The iterative process combined experts' knowledge and opinions to develop a group consensus of opinion (McKenna, 1994). The feedback provided by participants informed further development of the BATS.

In rounds one and two, participants were asked to consider items selected for potential inclusion in the BATS and rate the extent to which they agreed:

- Items were important and should be included in the BATS
- Items were comprehensible (see Figure 1).

In round three, item importance only was assessed. Participants were invited to use free text spaces to provide comments about the items, including suggestions for rewording.

| | (1) strongly disagree | (2) disagree | (3) disagree somewhat | (4) neither disagree or agree | (5) agree somewhat | (6) agree | (7) strongly agree |
|---|---|---|---|---|---|---|---|
| The item is important | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| The item is comprehensible | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

**Figure 1: Screenshot from the first round questionnaire of the Likert-type scales used to rate item importance and competence**

| Inclusion criteria | Exclusion criteria |
|---|---|
| **Article selection**: | |
| 1. Described the development and/or validation of a treatment fidelity measure. | Did not describe the development or validation of a treatment fidelity measure. |
| 2. Published in a peer reviewed journal. | Conference, dissertation, and thesis abstracts. |
| **Measure selection**: | |
| 3. Assessed treatment adherence and/or therapist competence. | Did not assess treatment adherence or therapist competence. |
| 4. Assessed therapist behaviours. | Assessed only patient behaviours. |
| 5. Evaluated the delivery of psychological therapies. | Did not evaluate the delivery of psychological therapies, including programme fidelity[1]. |
| 6. Evaluated therapies typically used for addressing alcohol and drug use problems, including those developed in other clinical areas. | Evaluated therapies not widely used in addictions, and not tailored to address alcohol or drug use problems. |
| 6b. Evaluated therapies not typically used in addictions but were tailored to address alcohol or drug use problems. | |
| 7. Evaluated individual therapies. | Evaluated couples, group, or family therapies. |
| 8. Evaluated therapies delivered face-to-face. | Evaluated therapies delivered online or by telephone. |
| 9. Evaluated therapies delivered in the home or in healthcare settings. | Evaluated therapies that are not delivered in the home or in healthcare settings. |
| 10. Target adult populations (16 years or over). | Target populations under the age of 16 years. |
| 11. Written in the English Language. | Not written in the English Language. |

[1]Programme fidelity covers multiple interventions and procedures (Bond et al., 2000) of which the delivery of a psychological therapy may be one aspect.

**Table 1: Criteria for article and measure selection**

| Phase | Overview of what the phase involved |
|---|---|
| Familiarisation with the data | The identified measures, particularly the items, were read and re-read. Initial analytic thoughts were noted. |
| Coding the data | The items from the identified measures were coded (*n*=833). This involved generating labels for identifying salient features of the data. |
| Searching for themes | Codes were grouped together to form themes, and sub-themes (lower level themes). Coded items were collated for each theme. |
| Reviewing themes | Themes were reviewed and refined to better reflect the coded items, and the dataset as a whole. |
| Defining themes | Definitions and labels for each theme were generated. A thematic map was developed. |
| Writing up | The analysis was written up to provide a coherent and plausible account of the data. The written narrative was supported by examples of the coded items. |

**Table 2: Phases of the thematic analysis**

## Consultation with current practitioners

10 practitioners and practice supervisors were consulted on presentation of the scale, its background and development to date at an addiction service peer supervision meeting. These therapists then used the scale to rate a video recorded practice session and provided feedback. Following discussion of this feedback in the research group, minor amendments to the item definitions were made.

## Using the scale to rate therapy sessions

The BATS was used to rate therapy sessions, which had been recorded as part of routine practice by therapists working at an addiction service, and three randomised controlled trials: UKATT (UKATT Research Team, 2005), AESOPS (Watson et al., 2013a), and ADAPTA (Watson et al., 2013b).

For the purposes of testing the reliability and validity of the scale, tapes were randomly selected from the four samples (see Table 3). Eighty recordings were randomly selected for independent process rating, 20 of which were double rated (see Figure 2). Raters (HC for independent rating, HC and GT for double rating) listened to each recording and used the BATS to score the extent to which therapists carried out the item behaviours. Independent rating enabled an exploration of differences in scores between therapies, and the testing of concurrent validity. Concurrent validity was explored by comparing the total scores on the BATS with process rating scores from ADAPTA, AESOPS, and UKATT, using Pearson's rho correlation coefficients (Field, 2009). Double rating enabled inter-rater reliability testing. For the purpose of this report, inter-rater reliability of the items was examined using the Intraclass Correlation Coefficient (ICC) two-way mixed-effects model (3.1) (Shrout and Fleiss, 1979) to aid comparison with previous literature.

| Type | Source | Data |
|---|---|---|
| Secondary analysis of trial data | ADAPTA (Watson et al., 2015) | Video recordings of 30-45 minute sessions of an alcohol focused (AF), and a healthy living (HL) intervention (*n*=50). |
| | AESOPS (Watson et al., 2013a) | Audio recordings of a 5-minute session of brief advice (BA), and a 20-minute session of behaviour change counselling (BCC) (*n*=160). |
| | UKATT (UKATT Research Team, 2005) | Video recordings of 50-minute sessions of motivational enhancement therapy (MET), and social behaviour and network therapy (SBNT) (*n*=452). |
| Primary routine practice Data | NHS specialist addictions service (SAS) | Video recordings of therapy sessions delivered by therapists working in a SAS in England (*n*=16). |
| | Non-NHS drug and alcohol service (DAS) | Video recordings of therapy sessions delivered by therapists working in a DAS in Wales (*n*=9). |

## Table 3: Data sources used in the validation of the BATS

**ADAPTA** = Addressing Drinking Among Patients: comparing Two Approaches (An alcohol-focused intervention versus a healthy living intervention for problem drinkers identified in a general hospital setting); **AESOPS** = Alcohol: Evaluating Stepped care in Older Populations Study; **UKATT** = United Kingdom (UK) Alcohol Treatment Trial.

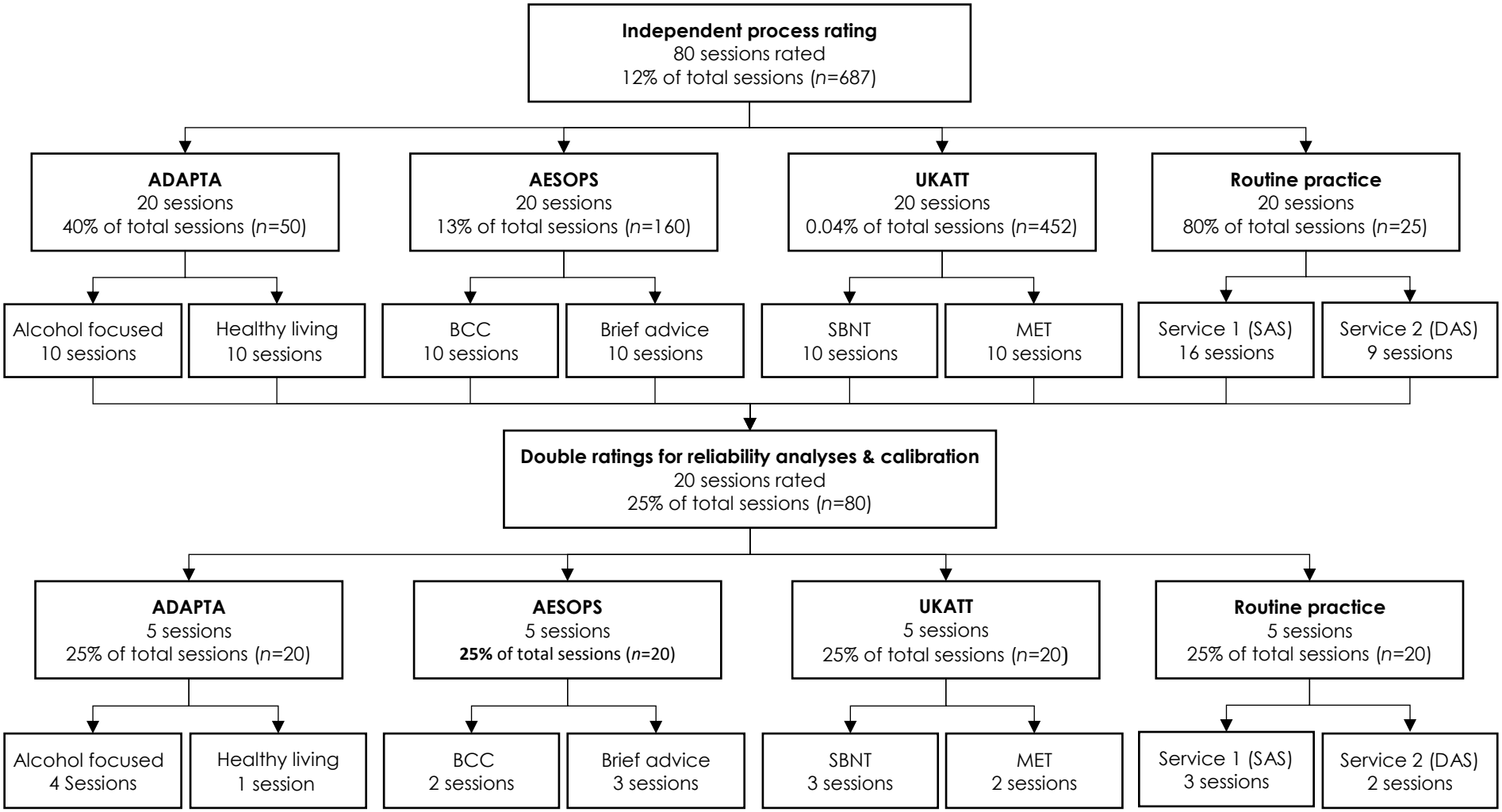



**Figure 2: Flow diagram of the selection of recordings for independent and double rating**

# FINDINGS

## Generating an item pool

The literature review identified 26 fidelity measures that evaluated the delivery of psychological therapies for addressing alcohol and drug use problems (see Figure 3). The measures included 783 items, describing the activities of therapists, including techniques associated with psychological interventions.

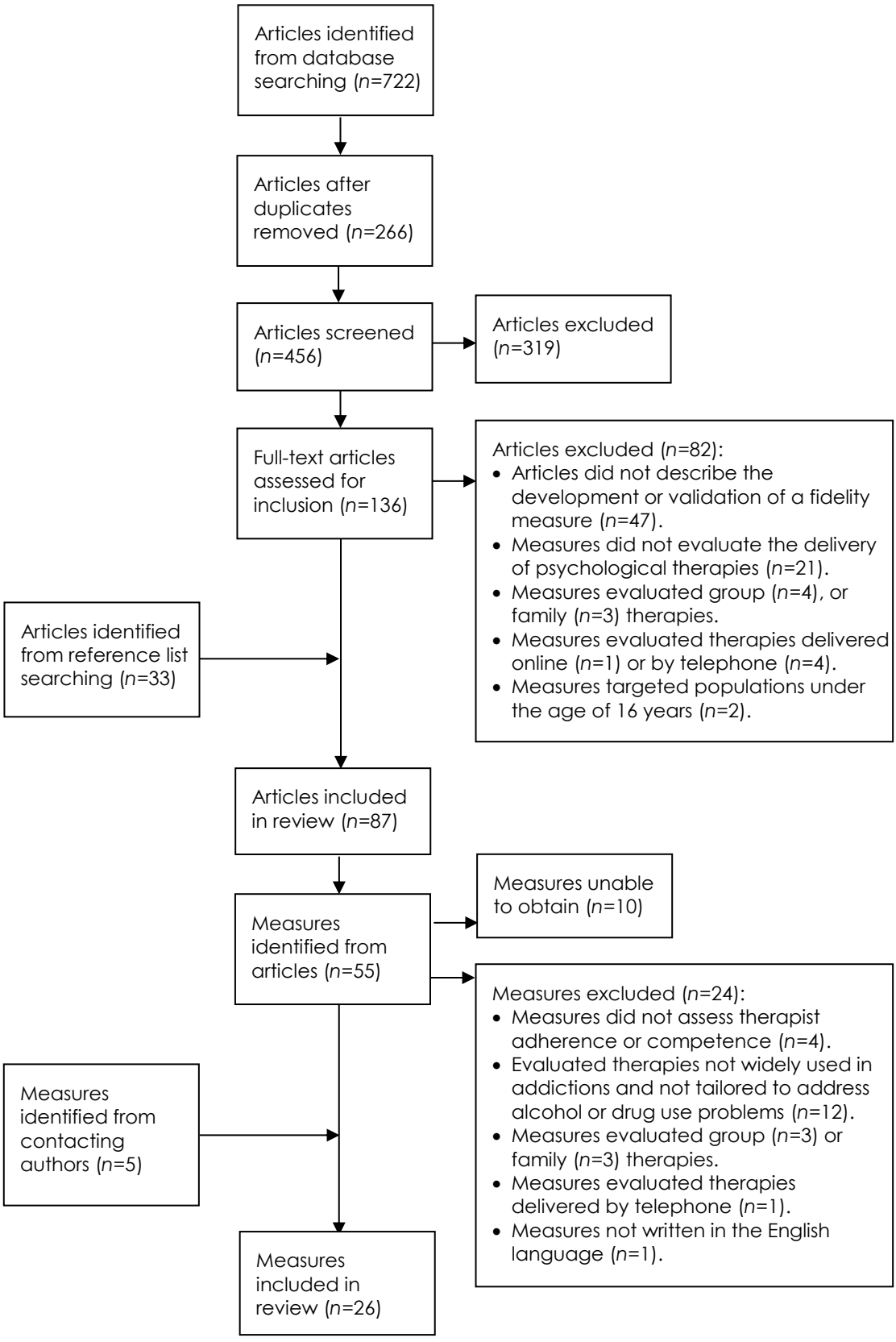


**Figure 3: Flowchart illustrating the selection of measures included in the review**

Thematic analysis of the items identified 18 themes (see Figure 4). Themes were grouped into four meta-themes:

- *Meta-theme 1: session management*. This related to therapists' management of the session, that is, keeping the session focused on the aims for that session.
- *Meta-theme 2: interventions to increase awareness*. This focused on techniques therapists may use to increase clients' understanding of their behaviours, thoughts, feelings and relationships, for example, by exploring clients' conflicting thoughts about changing their behaviour.
- *Meta-theme 3: interventions to change behaviour*. This considered techniques therapists may use to help clients change their behaviour and achieve their treatment goals, for example, by involving others, or encouraging talk about behaviour change.
- *Meta-theme 4: core skills*. This related to how therapists delivered the session, for example by developing an empathic and collaborative relationship (see Figure 4).

Exemplar items were developed from each theme. Exemplars were items that best reflected the theme descriptions. Most exemplars were adapted from existing fidelity measures ($n$=10), eight exemplars were newly constructed. The 18 exemplars were included in round one of the Delphi survey.

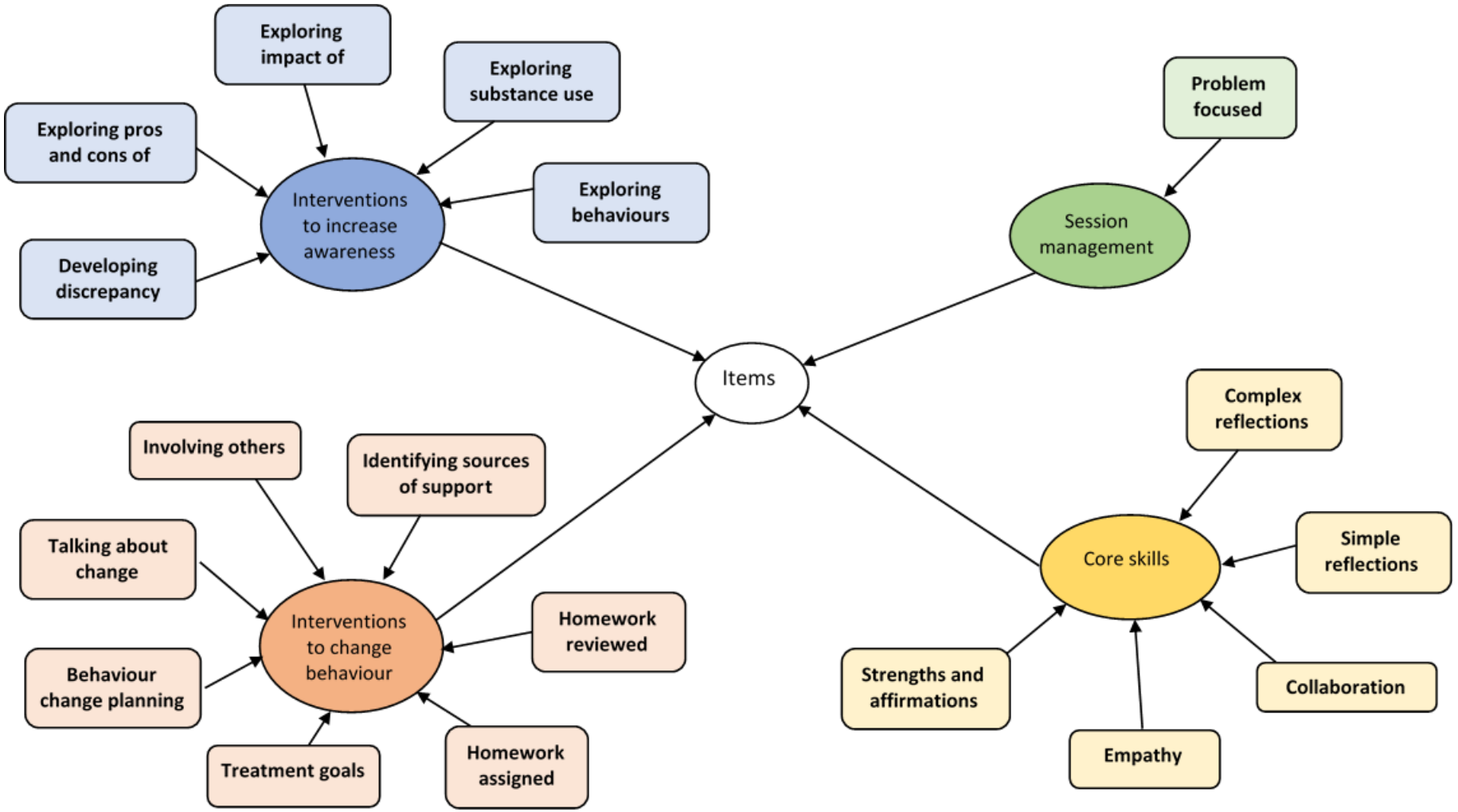



**Figure 4: Thematic map showing the four meta-themes and associated themes**

## Refining the scale content

Of the 19 experts invited to take part in the Delphi survey, 12 participants took part in round one, 12 in round two, and 10 in round three. Participants came from Europe and America, and all had over 15 years experience in the areas of addiction and psychotherapy. Most had both clinical and academic components to their work role. The diversity of participants' backgrounds assured a wide base of knowledge and expertise (Powell, 2003).

The results of the Delphi survey are summarised in Figure 5. The aim of round one was to reduce the number of items and to improve item comprehensibility. Round two aimed to reduce the level of dispersion among participants' views and further improve item comprehensibility. Round three focused on item importance. The aim of round three was to reach a consensus on the items to include in the BATS. The results showed group agreement on the scale content, and the final list of 12 items was included in the BATS (see Appendix 1).

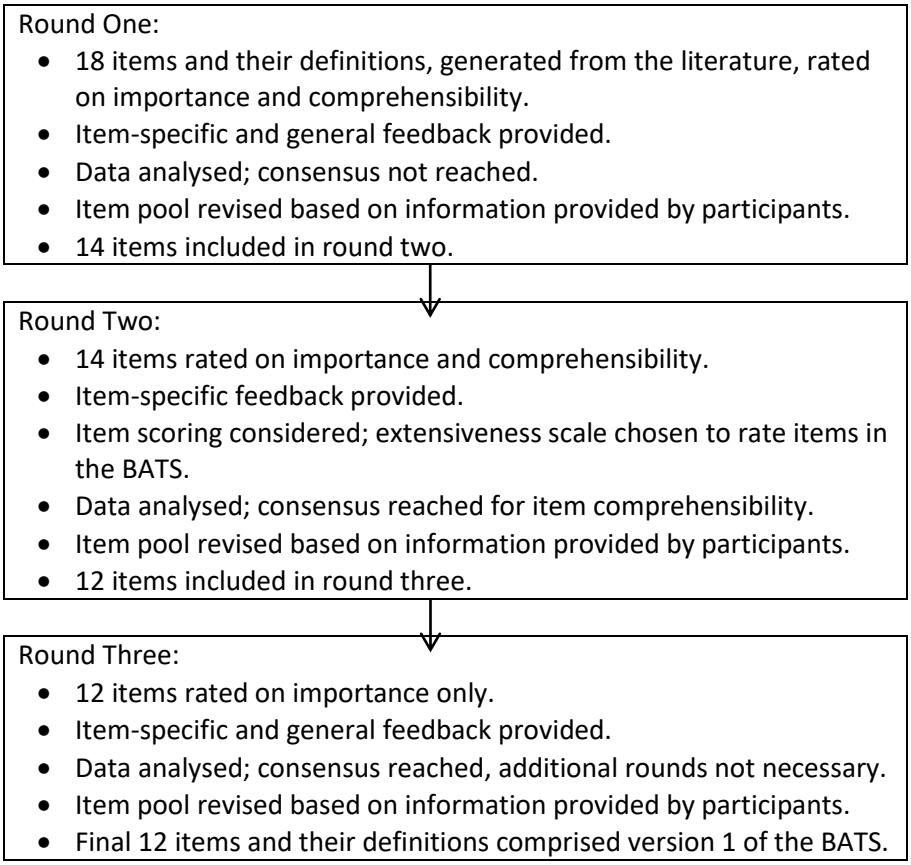


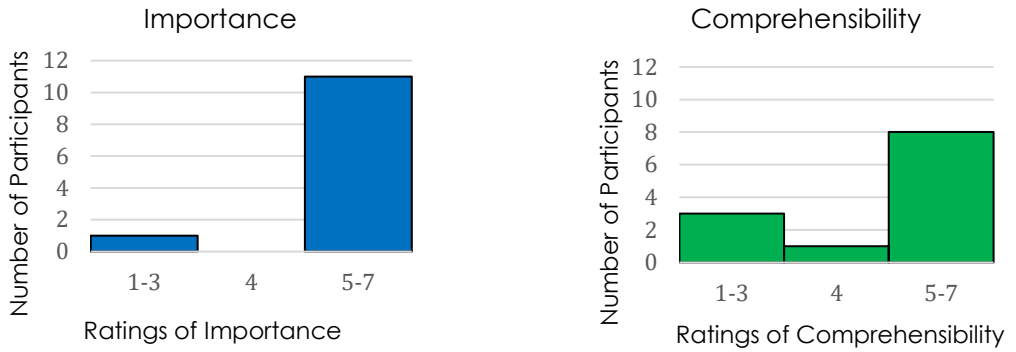

**Figure 5: Summary of the Delphi survey**

Figure 6 illustrates how consensus was reached for item 1 'problem focused'; a shift towards agreement can be seen for both importance and comprehensibility when the second and third rounds are compared to the first round. Based on the results of the Delphi survey, version 1 of the BATS was created with guidance notes to aid completion.

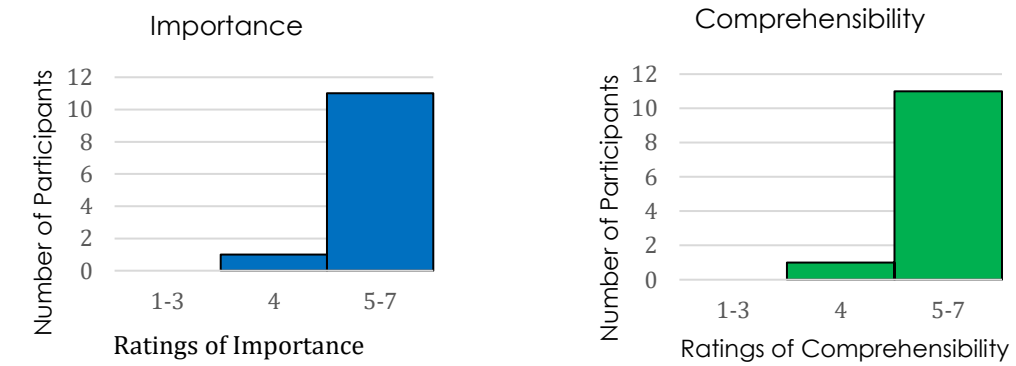
## Using the scale to rate therapy sessions

The average length of the session recordings ranged from 6 minutes (BA sessions) to 55 minutes (SBNT sessions). Differences in item scores can be seen in Table 4. Concurrent validity was demonstrated by significant correlations with the ADAPTA process rating scale (Pearson's $r$ = 0.678, $p<0.01$, n=20); and the AESOPS process rating scale (Pearson's $r$ = 0.805, $p<0.001$ n=20).

***Round one***: Frequency distributions of participants' importance and comprehensiblity ratings for item 1




***Round two***: Frequency distributions of participants' importance and comprehensiblity ratings for item 1




***Round three***: Frequency distributions of participants' importance ratings for item 1

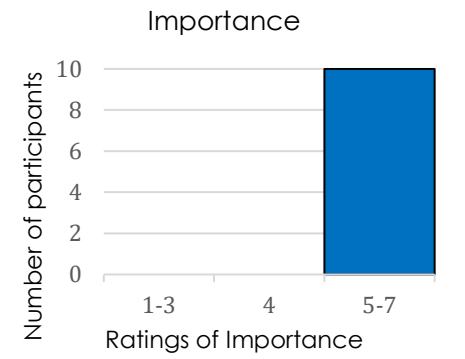



**Figure 6: Illustration of how consensus was reached for item 1 'problem focused'**

| Item reference | | Therapy | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | AF (*n*=10) | HL (*n*=10) | BA (*n*=10) | BCC (*n*=10) | MET (*n*=10) | SBNT (*n*=10) | RP (*n*=20) | Total (*n*=80) |
| | | Median* (IQR), range | Median* (IQR), range | Median* (IQR), range | Median* (IQR), range | Median* (IQR), range | Median* (IQR), range | Median* (IQR), range | Median* (IQR), range |
| 1. | Problem focused | 4 (0), 3-4 | 3.5 (2), 1-4 | 3.5 (1), 1-4 | 3 (2), 2-4 | 3 (2), 1-4 | 2 (2), 1-4 | 4 (1), 1-4 | 3 (1), 1-4 |
| 2. | Collaboration | 3.5 (1), 3-4 | 4 (1), 3-4 | 1 (1), 0-2 | 2 (2), 1-4 | 2 (3), 0-4 | 1 (2), 0-4 | 4 (1), 2-4 | 3 (2), 0-4 |
| 3. | Empathy | 3 (1), 2-4 | 4 (1), 3-4) | 2 (1), 1-3 | 2.5 (1), 2-4 | 2 (1), 1-4 | 2 (2), 1-4 | 4 (1), 2-4 | 3 (2), 0-4 |
| 4. | Strengths and affirmation | 1.5 (3), 0-4 | 1 (2), 0-4 | 0 (0), 0-1 | 1 (1), 0-2 | 1 (2), 0-3 | 1 (3), 0-3 | 2 (2), 0-4 | 1 (2), 0-4 |
| 5. | Complex reflections | 3 (2), 2-4 | 3 (1), 2-4 | 0 (1), 0-1 | 2 (1), 1-3 | 2 (2), 1-4 | 1 (1), 1-4 | 2 (3), 0-4 | 2 (2), 0-4 |
| 6. | Homework assigned | 2 (2), 0-3 | 1 (1), 1-4 | 0 (0), 0-0 | 0 (0), 0-0 | 0 (0), 0-1 | 0.5 (0), 0-1 | 1.5 (3), 0-4 | 0 (2), 0-4 |
| 7. | Homework reviewed | 1 (1), 0-3 | 0.5 (3), 0-4 | 0 (0), 0-0 | 0 (0), 0-0 | 0 (0), 0-3 | 0 (0), 0-1 | 1 (1), 0-4 | 0 (2), 0-4 |
| 8. | Treatment goals | 3.5 (1), 1-4 | 2 (2), 1-4 | 0 (1), 0-1 | 2.5 (2), 0-4 | 2 (2), 1-4 | 2 (2), 0-4 | 3 (2), 0-4 | 2 (2), 0-4 |
| 9. | Developing discrepancy | 0 (1), 0-2 | 0 (0), 0-1 | 0 (0), 0-0 | 0 (0), 0-1 | 1 (2), 0-2 | 0 (1), 0-1 | 0 (0), 0-3 | 0 (1), 0-3 |
| 10. | Exploring pros and cons of change | 1 (2), 0-2 | 1 (2), 0-3 | 0 (0), 0-1 | 1 (1), 0-2 | 1.5 (3), 0-4 | 0 (0), 0-1 | 1 (2), 0-4 | 1 (2), 0-4 |
| 11. | Behaviour change planning | 1 (2), 0-3 | 2 (2), 1-4 | 0 (0), 0-1 | 1 (1), 0-3 | 0.5 (2), 0-3 | 1.5 (2), 1-4 | 2 (2), 0-4 | 1 (3), 0-4 |
| 12. | Sources of support | 3 (0), 2-4 | 1 (1), 0-4 | 0 (0), 0-1 | 0 (0), 0-0 | 0.5 (1), 0-2 | 2 (3), 1-4 | 1 (1), 0-4 | 1 (2), 0-4 |

*Ratings made on a 5-point Likert scale: 0 = not at all, 1 = a little, 2 = somewhat, 3 = a good deal, 4 = extensively; IQR = interquartile range; AF = alcohol focused intervention (ADAPTA); HL = health living focused intervention (ADAPTA); BA = brief advice (AESOPS); BCC = behaviour change counselling (AESOPS); MET = motivational enhancement therapy (UKATT); SBNT = social behaviour and network therapy (UKATT); RP = routine practice therapy sessions.

**Table 4: Statistical summary of the scores for the rated therapy sessions**

A sample of the recordings was rated by two raters (HC & GT) (*n*=20). ICCs for the items ranged from 0.76 to 0.96, indicating excellent reliability (see Table 5).

| Item Reference | | Mean Ratings* (SD) | | ICC (95% CI) |
|---|---|---|---|---|
| | | Rater 1 | Rater 2 | |
| 1. | Problem focused | 3.40 (1.00) | 3.25 (1.25) | 0.87 (0.69 to 0.94) |
| 2. | Collaboration | 2.90 (1.37) | 2.95 (1.43) | 0.93 (0.84 to 0.97) |
| 3. | Empathy | 3.00 (1.12) | 2.65 (1.39) | 0.93 (0.82 to 0.97) |
| 4. | Strengths and affirmation | 2.05 (1.47) | 2.15 (1.60) | 0.96 (0.90 to 0.98) |
| 5. | Complex reflections | 2.15 (1.35) | 1.60 (1.39) | 0.85 (0.65 to 0.94) |
| 6. | Homework assigned | 0.80 (1.29) | 1.00 (1.26) | 0.92 (0.80 to 0.97) |
| 7. | Homework reviewed | 0.55 (1.00) | 0.75 (1.16) | 0.88 (0.73 to 0.95) |
| 8. | Treatment goals | 2.45 (1.47) | 2.30 (1.26) | 0.74 (0.45 to 0.90) |
| 9. | Developing discrepancy | 0.50 (0.69) | 0.75 (1.02) | 0.87 (0.70 to 0.95) |
| 10. | Exploring pros and cons of change | 1.25 (1.16) | 1.30 (1.13) | 0.82 (0.60 to 0.93) |
| 11. | Behaviour change planning | 1.55 (1.36) | 1.80 (1.40) | 0.81 (0.58 to 0.92) |
| 12. | Sources of support | 1.50 (1.24) | 3.25 (1.25) | 0.87 (0.69 to 0.94) |

*Ratings made on a 5-point Likert scale: 0 = not at all, 1 = a little, 2 = somewhat, 3 = a good deal, 4 = extensively. Mean ratings based on 20 recordings.
SD = standard deviation; ICC = intraclass correlation coefficient; CI = confidence interval

**Table 5: Intraclass correlation coefficient analyses of the individual items**

# DISCUSSION

The addiction treatment field has been inundated with different approaches to understanding and treating substance misuse and addiction problems. In recent years, we have seen concerted efforts to identify effectiveness findings from among the myriad approaches. In this project, the first task was to distil out the essential 'active ingredients' of treatment and to combine these into a rating scale that could be used to guide and to monitor practice in routine settings. The development of the BATS was based upon consensus about core active ingredients of effective practice both from the literature and from expert consultation. The challenge then was to operationalize these into a valid and reliable scale.

The next task is further to test acceptability and usability by making the scale widely available and encouraging its use. Ensuring inter-rater reliability beyond the research team is a key element of this task. In a climate of target-setting and performance management, which is based upon quantity rather than quality, attention to standards of practice that are research-based (and relatively easy to implement within organisational constraints) is essential.

## Study Impact

The BATS has been incorporated into routine practice at an NHS specialist addiction service to support peer supervision. Permission to use the BATS has been granted to addiction services in Estonia and Wales. The Leeds Addiction Therapy Unit have plans to support its implementation. The study has been presented at three local and one international conference.

## Conclusion

This study has developed a brief, evidence-based tool for monitoring and evaluating the delivery of psychological therapies used in routine practice. The BATS is trans-theoretical, applicable to the range of widely-used effective approaches. It has good face validity, and good to excellent inter-rater reliability. Users have commented that the guidance is brief and easy to use. That there is no requirement for a lengthy manual is important if the scale is to be implemented in routine practice.

## Next Steps for the BATS

The psychometric properties of the BATS continue to be investigated through feedback from use and future data collection at the pilot sites. Semi-structured interviews with therapists are also being undertaken to assess usability and ensure terminology is meaningful to the target users of the scale. The BATS will be incorporated in an addictions assessment and outcome measurement website www.result4addiction.net.

# REFERENCES


Aarons, G.A., Hurlburt, M. and Horwitz, S.M. 2011. Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health and Mental Health Services Research.* 38(1), pp.4-23.

Braun, V. and Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology.* 3(2), pp.77-101.

Carroll, K.M. 1998. *A cognitive behavioral approach: treating cocaine addiction.* National Institute on Drug Abuse Rockville, MD.

Carroll, K.M., Nich, C. and Rounsaville, B.J. 1998. Utility of therapist session checklists to monitor delivery of coping skills treatment for cocaine abusers. *Psychotherapy Research.* 8(3), pp.307-320.

Carroll, K.M., Nich, C., Sifry, R.L., Nuro, K.F., Frankforter, T.L., Ball, S.A., Fenton, L. and Rounsaville, B.J. 2000. A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug and Alcohol Dependence.* 57(3), pp.225-238.

Copello, A., Orford, J., Hodgson, R., Tober, G., Barrett, C. and Team, U.R. 2002. Social behaviour and network therapy: basic principles and early experiences. *Addictive behaviors.* 27(3), pp.345-366.

Field, A. 2009. *Discovering Statistics Using IBM SPSS Statistics (Third Edition).* London: SAGE.

Hill, C.E., O'Grady, K.E. and Elkin, I. 1992. Applying the collaborative study psychotherapy rating scale to rate therapist adherence in cognitive-behavior therapy, interpersonal therapy, and clinical management. *Journal of Consulting and Clinical Psychology.* 60(1), pp.73-79.

Madson, M.B. and Campbell, T.C. 2006. Measures of fidelity in motivational enhancement: A systematic review. *Journal of Substance Abuse Treatment.* 31(1), pp.67-73.

Madson, M.B., Campbell, T.C., Barrett, D.E., Brondino, M.J. and Melchert, T.P. 2005. Development of the Motivational Interviewing Supervision and Training Scale. *Psychology of Addictive Behaviors.* 19(3), pp.303-310.

Martino, S., Ball, S.A., Nich, C., Frankforter, T.L. and Carroll, K.M. 2008. Community program therapist adherence and competence in motivational enhancement therapy. *Drug and Alcohol Dependence.* 96(1–2), pp.37-48.

McKenna, H.P. 1994. The Delphi technique: a worthwhile research approach for nursing? *Journal of Advanced Nursing.* 19(6), pp.1221-1225.

Middleton, W., Tober, G., Frier, E. and Finnegan, O., with the UKATT Process Data Sub-Group. 2001. *UKATT Process Rating Manual*. Unpublished.

Miller, W.R., Zweben, A., DiClemente, C.C. and Rychtarik, R. 1992. *Motivational Enhancement Therapy Manual: A Clinical Research Guide for Therapists Treating Individuals with Alcohol Abuse and Dependence. Project MATCH Monograph Series, Volume 2.* Rockville, MD: National Institute on Alcohol Abuse and Alcoholism.

Norcross, J.C. and Wampold, B.E. 2011. What works for whom: Tailoring psychotherapy to the person. *Journal of Clinical Psychology.* 67(2), pp.127-132.

Perepletchikova, F. and Kazdin, A.E. 2005. Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice.* 12(4), pp.365-383.

Powell, C. 2003. The Delphi technique: myths and realities. *Journal of Advanced Nursing.* 41(4), pp.376-382.

Project MATCH Research Group. 1997. Matching Alcoholism Treatments to Client Heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol.* 58(1), pp.7-29.

Raistrick, D., Heather, N. and Godfrey, C. 2006. *Review of the effectiveness of treatment for alcohol problems.* London.

Schoenwald, S.K. and Garland, A.F. 2013. A review of treatment adherence measurement methods. *Psychological assessment.* 25(1), p146.

Schoenwald, S.K., Garland, A.F., Chapman, J.E., Frazier, S.L., Sheidow, A.J. and Southam-Gerow, M.A. 2011. Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health and Mental Health Services Research.* 38(1), pp.32-43.

Shrout, P.E. and Fleiss, J.L. 1979. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin.* 86(2), pp.420-428.

Tober, G. and Crosby, H.F. 2011. *AESOPS Process Rating Manual Version 5*. Unpublished.

Tober, G. and Crosby, H.F. 2014. *ADAPTA Process Rating Scale Manual*. Unpublished.

Tober G, Clyne W, Finnegan O, Farrin A & Russell I in collaboration with the UKATT Research Team (2008) Validation of a scale for rating the delivery of psycho-social treatments for alcohol dependence and misuse: the UKATT Process Rating Scale (PRS). *Alcohol and Alcoholism* 43: 675-682.

UKATT Research Team. 2005. Effectiveness of treatment for alcohol problems: findings of the randomised UK alcohol treatment trial (UKATT). *British Medical Journal.* [Online].

331(7516), pp.1-5. [Accessed 08/02/2018]. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1200586/

Waltz, J., Addis, M.E., Koerner, K. and Jacobson, N.S. 1993. Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology*. 61(4), pp.620-630.

Watson, J., Crosby, H., Dale, V., Tober, G., Wu, Q., Lang, J., McGovern, R., Newbury-Birch, D., Parrott, S. and Bland, J. 2013a. AESOPS: a randomised controlled trial of the clinical effectiveness and cost-effectiveness of opportunistic screening and stepped care interventions for older hazardous alcohol users in primary care.

Watson, J., Tober, G., Raistrick, D., Mdege, N., Dale, V., Crosby, H., Godfrey, C., Lloyd, C., Toner, P. and Parrott, S. 2013b. An alcohol-focused intervention versus a healthy living intervention for problem drinkers identified in a general hospital setting (ADAPTA): study protocol for a randomized, controlled pilot trial. *Trials.* [Online]. 14(1), p117. [Accessed 06/07/2017]. Available from: https://trialsjournal.biomedcentral.com/track/pdf/10.1186/1745-6215-14-117?site=trialsjournal.biomedcentral.com

Watson, J.M., Fairhurst, C., Li, J., Tober, G., Crosby, H., Lloyd, C., Godfrey, C., Mdege, N.D., Dale, V., Toner, P., Parrott, S. and Raistrick, D. 2015. ADAPTA: A pilot randomised controlled trial of an alcohol-focused intervention versus a healthy living intervention for problem drinkers identified in a general hospital setting. *Drug and Alcohol Dependence*. 154, pp.117-124.

## APPENDIX 1: BATS version 1

**B A T S**

**Brief Addiction Therapist Scale**

**A tool for evaluating therapists' delivery of psychological therapies for alcohol and drug use problems.**

**Designed to facilitate training and supervision, and enhance therapist skill.**

**Instructions for use**: For each item, circle a number on the 5-point scale reflecting the extent to which the therapist carried out the behaviour. For items that are not applicable to the session, score 0 'not at all'. Use the space provided on page 2 to give context, comments, and additional information e.g. the client's first session. Item definitions are provided on page 2. To be used with audio or video recordings of therapy sessions.

| **During the session…** | *Not at all* | *A little* | *Somewhat* | *A good deal* | *Extensively* |
|---|---|---|---|---|---|
| 1. The therapist kept the session focused on the aims for that session. | 0 | 1 | 2 | 3 | 4 |
| 2. The therapist attempted to work together with the client. | 0 | 1 | 2 | 3 | 4 |
| 3. The therapist conveyed empathy. | 0 | 1 | 2 | 3 | 4 |
| 4. The therapist focused on the client's strengths. | 0 | 1 | 2 | 3 | 4 |
| 5. The therapist used "complex reflections" – offering a perspective which added meaning and enabled the client to make connections. | 0 | 1 | 2 | 3 | 4 |
| 6. The therapist and the client planned tasks for the client to do between sessions. | 0 | 1 | 2 | 3 | 4 |
| 7. The therapist and the client reviewed tasks planned in the previous session. | 0 | 1 | 2 | 3 | 4 |
| 8. The therapist enabled the client's goals for treatment to be discussed. | 0 | 1 | 2 | 3 | 4 |
| **If in this session the focus was on building motivation for change:** | | | | | |
| 9. The therapist encouraged the client to consider inconsistencies between their substance use, and personal goals or values. | 0 | 1 | 2 | 3 | 4 |
| 10. The therapist encouraged the client to talk about the positive aspects of changing substance use. | 0 | 1 | 2 | 3 | 4 |
| **If in this session the focus was on planning or maintaining change:** | | | | | |
| 11. The therapist enabled a plan for changing the client's substance use, or maintaining change, to be discussed. | 0 | 1 | 2 | 3 | 4 |
| 12. The therapist discussed how the client's social network might support changing substance use or maintaining change. | 0 | 1 | 2 | 3 | 4 |

| **Total score:** | |
|---|---|

## Item definitions

1. **Session aims:** The therapist kept the session focused on clinically relevant aims during the session, e.g. target behaviour. This may or may not include explicit discussion of the purpose of the session, *e.g. to describe a relapse prevention plan*. Aims may change during the course of the session following disclosure of risk.
2. **Working together**: Developing a collaborative relationship between the client and the therapist. It is about discussing, actively seeking the client's input; not telling, and not arguing.
3. **Convey empathy**: Making efforts to convey warmth and understanding of the client's thoughts and feelings. The therapist avoids any blaming or labelling.
4. **The client's strengths**: Helping the client to identify and focus on what they can do, not what they cannot do: achievements rather than failings.
5. **Complex reflections**: Helping the client to gain insight by making and/or strengthening connections between things they have said. Going beyond repeating or slightly rephrasing what the client has said.
6. **Planning tasks**: Any task that is planned (the therapist and the client agreed what to do and how to do it) for the client to do between sessions, *e.g. specific homework tasks, trying new behaviours*.
7. **Reviewing tasks**: Explicit discussion in which tasks set in the previous session are reviewed. This item is not applicable if it is the client's first session, tick the box as appropriate.
8. **Treatment goals**: Goals refer to the overall treatment goals, *e.g. abstinence, harm reduction, moderation*. The goals could be discussed by the therapist and/or the client.
9. **Considering inconsistencies**: Exploring how the client's behaviour conflicts with his/her personal goals and values, *e.g. I need to drink a bottle of gin but I want to be a good parent*.
10. **Talking about change**: The therapist encourages the client to talk about the positive aspects of changing.
11. **Change planning**: Discussion of an overall plan to achieve the agreed treatment goals. Tasks represent the steps in the plan to achieve the overall treatment goals.
12. **The social network**: The therapist facilitates a discussion about the client's actual and/or potential social network to identify how this may support the overall plan.

## Context, comments, and additional information:


We acknowledge the support of Leeds and York Partnership NHS Foundation Trust (LYPFT). BATS was funded by Alcohol Research UK (ARUK), and the Society for the Study of Addiction (SSA). We thank the participants involved in developing BATS.

UNIVERSITY OF LEEDS