

## Evaluation of the ACTivate Your Life Course

Dr Evelyn Gibson, Principal Clinical Psychologist, Lead for Governance / Delivery Outcomes, Research & Audit, Hywel Dda University Health Board

**Abstract:** *ACTivate Your Life (AYL) is an evidence-based, group intervention for mild to moderate anxiety and depression. Hywel Dda University Health Board (H DUHB) has been running this course for a number of years. Study 1 aimed to determine the impact of this intervention. Participants were asked to complete session-by-session measures for anxiety, depression, general psychological distress and psychological flexibility. Participants and group facilitators also provided qualitative feedback on their experience of the course. The overall group results indicated that there was a statistically significant change on all outcome measures. On an individual level, reliable recovery rates were between 8% and 42%. Study 2 took historical pre- and post-intervention data collected in one county and used this to determine whether there had been any change to the efficacy of the course. The results indicated that the outcomes for this intervention were less positive than those found in Study 1; however, participants rated the intervention in Study 2 as better than that in Study 1. Recommendations for the future of AYL course in H DUHB are presented.*

### Introduction

The Welsh model for the provision of mental health services defines a number of ‘tiers’ of increasing intensity of intervention (Welsh Government, 2012). Tier 0, the least intensive tier, involves high volume and low intensity interventions, such as group-based psychoeducation courses. This is in-line with the suggestion by the National Institute for Health and Care Excellence (NICE, 2011) that psychoeducation should be considered before the use of complex assessment and treatment strategies for individuals presenting with mild symptoms. Psychoeducation has been shown to be effective in treating a range mental health disorders (Lukens & McFarlane, 2004).

One such psychoeducational group intervention is “*ACTivate Your Life*” (AYL), a four-week programme that is offered throughout Wales. AYL is a transdiagnostic intervention for adults with mild to moderate mental health conditions that is based on Acceptance and Commitment Therapy (ACT). AYL offers an alternative approach to equivalent groups based on a more traditional cognitive behavioural therapy (CBT) approach. Rather than focusing on changing psychological events directly, ACT seeks to change the function of events and the individual’s relationship to them. There are also versions of the course that have been adapted for specific groups including university students, stroke survivors, individuals with chronic pain and individuals who lives have been affected by cancer.

The ACT model postulates that attempts to change aversive internal stimuli may increase distress (Hayes & Strosahl, 2004). ACT treatment consists of developing awareness and non-judgmental acceptance of both negative and positive experiences; identification of valued life directions; and, making appropriate action towards goals that support those values (Hayes, Strosahl & Wilson, 1999). The overarching aim of ACT is the development of psychological flexibility. The ACT approach can be defined in terms of six key psychological processes:

1. *Experiential avoidance/acceptance* - Experiential avoidance refers to efforts to alter the frequency or form of unwanted private events (e.g., thoughts, memories, emotions, bodily sensations) even when doing so causes distress (Hayes, Wilson, Gifford, Follette & Strosahl, 1996). It is associated with a wide variety of negative mental health outcomes, including high

levels of anxiety, depression and psychosocial difficulties (Hayes, Luoma, Bond, Masuda & Lillis, 2006). ACT focuses on developing acceptance as an alternative to experiential avoidance. Acceptance is defined as the active awareness of private experiences without attempts to change them.

2. *Cognitive fusion/defusion* - Cognitive fusion refers to the tendency to become caught up in thoughts due to a strong belief in their literal content. This can cause behaviour to become narrow, rigid and less guided by experience. Cognitive fusion has been shown to contribute negatively to chronic pain (Wicksell, Renofalt, Olsson, Bond & Melin, 2008); mental health problems in children and adolescents (Greco, Lambert & Baer, 2008); and, depression in adults (Addis & Jacobson, 1996). ACT uses cognitive defusion and mindfulness techniques to create more flexibility in the presence of difficult thoughts. Defusion strategies alter the functional context of cognitive events by decreasing their believability and impact. In contrast, CBT techniques attempt to reduce the frequency of the thought; challenge its validity; or, test it in the real world, thus treating the thought as if it is important.
3. *Being present* - Individuals often focus their attention on the past or the future. This is known to exacerbate problems such as trauma (Holman & Silver, 1998), rumination (Davis & Nolen-Hoeksema, 2000) and pain (Schutze, Rees, Preece & Schütze, 2010). ACT uses mindfulness and attentional control exercises to promote focused, voluntary and flexible contact with the present moment.
4. *Conceptualised self / noticing self* - When people are asked about themselves, they tend to describe their conceptualised self or self-narrative (e.g., "*I am someone who always tries hard*"). The conceptualised self often reduces behavioural flexibility. Moreover, events that threaten the conceptualised self can evoke strong emotions and lead to heightened experiential avoidance (Mendolia & Baker, 2008). Directly changing self-concepts can be difficult (Baumeister, Campbell, Krueger & Vohs, 2003), so ACT teaches ways to develop a noticing sense of self. This provides a secure psychological space for facing painful emotions or thoughts (Hayes, 1984).
5. *Unclear, compliant or avoidant motives/values* - When behaviour change is motivated by guilt or compliance, goal achievement is much less likely (e.g., Elliot, Sheldon & Church, 1997; Sheldon & Elliot, 1999; Sheldon, Kasser, Smith & Share, 2002). ACT seeks to link behaviour to client values which act as intrinsic reinforcers.
6. *Committed action* - ACT encourages the continuous redirection of behaviour to produce patterns of effective action linked to chosen values. ACT methods have been shown to foster higher levels of committed actions such as willingness to engage in exposure therapy (Levitt, Brown, Orsillo & Barlow, 2004) or to seek employment (Dahl, Wilson, Luciano & Hayes, 2005).

A review of the meta-analytic evidence on ACT, found that ACT was generally superior to inactive controls, treatment as usual and most active intervention conditions excluding CBT (Gloster, Walder, Levin, Twohig & Karekla, 2020). ACT has been found to be effective in depression (Zettle, Rains & Hayes, 2011), chronic pain (Wicksell, Ahlqvist, Bring, Melin & Olsson, 2008), coping with psychosis (Bach & Hayes, 2002), obsessive compulsive disorder (Twohig *et al.*, 2010), mixed anxiety disorders (Arch *et al.*, 2012) and substance misuse (Twohig, Shoenberger & Hayes, 2007; Gifford *et al.*, 2004). Group-based ACT has been applied to numerous physical health conditions and has been shown to reduce depression and anxiety and to improve condition-specific outcomes (Kemani *et al.*, 2015; McCracken, Sato & Taylor, 2013; Mohabbat-Bahar, Maleki-Rizi, Akbari & Moradi-Joo, 2015; Nordin & Rorsman, 2012).

AYL has been running in Hywel Dda University Health Board (HDUHB) for a number of years; however, due to staff turnover many of the facilitators that were originally trained in presenting AYL are no longer involved. It was felt that the AYL courses should be evaluated to ensure that the efficacy of this group has been maintained over time in light of the fact that there has been no ongoing AYL-

specific training. This evaluation is described in Study 1 and addresses the following research questions: (1) How effective is AYL in HDUHB? (2) Is the efficacy of the course similar to that found in other Health Boards in Wales and/or published studies? (3) Are clinical outcomes influenced by initial symptom severity? (4) What are participants' and facilitators' experiences of AYL?

Study 2 uses historical outcome data for AYL groups that have been run in Ceredigion over the past few years. This study addresses the following research questions: (1) How effective were the AYL that were run historically in HDUHB? (2) How do the outcomes from past AYL groups compare to the outcomes of the AYL groups that are currently being run in HDUHB?

## **STUDY ONE: THE EFFECTIVENESS OF THE CURRENT AYL GROUPS RUN IN HDUHB**

### **Method**

#### *Intervention*

Seven AYL groups were run across the HDUHB region between October 2019 and February 2020. A further group was started but had to be discontinued due to the UK-wide restrictions imposed in March 2020 related to the coronavirus pandemic. Available demographic and outcome data were aggregated for all individuals who accessed AYL across the HDUHB region for the analysis. Demographic data from the discontinued group was also included in the analysis.

Participants attended AYL through two routes (1) referred to Local Primary Care Mental Health Support Services (LPMHSS) by GPs; or, (2) self-referral through gaining knowledge of AYL online; via posters; through word of mouth; or, through other services.

Each AYL group was facilitated by LPMHSS staff. AYL ran weekly over four, two-hour sessions. Participants were not followed-up if they missed sessions and were not reviewed on completion of the course.

#### *Design*

A mixed methods design was employed. The strategy involved primary quantitative methods and secondary qualitative methods used to obtain the views of both course participants and facilitators<sup>1</sup>. The quantitative aspects of the analysis provided statistical and clinically relevant data regarding the outcomes of the intervention, while the qualitative aspects ensured that the complexity inherent in the experiences of participants and facilitators was not lost. Quantitative data was analysed using SPSS Statistics (IBM Corp., 2017).

A repeated measures design was used for the quantitative evaluation. Repeated measures ANOVAs were used to analyse changes in symptoms and overall distress levels for each of the four sessions. In addition, dependent t-tests on the results of psychometric measures were used to obtain overall pre- and post-intervention changes. Calculations of clinical significance and reliable recovery rates were also carried out. Descriptive statistics were produced for the demographic data.

Content analysis was employed to analyse the Initial and Evaluation Questionnaires returned by the participants. The broad categories of responses were refined to elucidate underlying meanings in an iterative and inductive process to generate higher-order themes.

---

<sup>1</sup> Due to time limitations, facilitator feedback was only obtained from LPMHSS staff in Carmarthenshire

## *Questionnaires*

### Initial and Evaluation Questionnaires

The Initial Questionnaire gathered demographic information; information related to the source of participants' recruitment; and, the goals that participants had at the start of the course. This questionnaire required categorical responses. The Evaluation Questionnaire gathered feedback on participants' experience of attending AYL. It posed open-ended questions on attendees' experiences of the intervention, alongside Likert scale questions to determine participants' views on the utility of the course.

### CORE-10 (Connell & Barkham, 2007)

The CORE-10 is a brief self-report questionnaire comprising of ten items that are designed to measure an individual's level of global distress, including commonly experienced symptoms of anxiety and depression and aspects of life and social functioning. Each item within the CORE-10 is scored on a five-point Likert scale.

The CORE-10 distinguishes between clinical and non-clinical populations. It has good specificity; acceptability; sensitivity to change; and, very good internal reliability (Barkham *et al.*, 2013).

### GAD-7 (Spitzer, Kroenke, Williams & Löwe, 2006)

The GAD-7 was developed as a self-report screening tool for generalised anxiety disorder (GAD); however, it is now used to measure symptomology for a variety of anxiety disorders (Beard & Björgvinsson, 2014). The GAD-7 consists of seven items each on a four-point Likert scale.

The GAD-7 is considered to be a valid and reliable case identifier and outcome measure for anxiety in both psychiatric samples (Kertz, Bigda-Peyton & Björgvinsson, 2013) and in the general population (Löwe *et al.*, 2008). It has good sensitivity and specificity; excellent internal consistency; and, good test-retest reliability (Spitzer *et al.*, 2006).

### PHQ-9 (Kroenke, Spitzer & Williams, 2001)

The PHQ-9 was originally developed as a self-report screening tool for depression in primary care. It consists of nine items, each on a four-point Likert scale.

The PHQ-9 is considered to be a valid and reliable case identifier and outcome measure for depression (Spitzer *et al.*, 2006). It has been reported to have good internal consistency; good sensitivity to change; and, good specificity (Beard, Hsu, Rifkin, Busch & Björgvinsson, 2016).

### The Acceptance and Action Questionnaire (AAQ-ii; 2nd version; Bond *et al.*, 2011)

The AAQ-ii is a measure of experiential avoidance and psychological inflexibility and, as such, can be used as a process measure for ACT-based interventions. The AAQ-ii has been found to predict a wide range of quality-of-life outcomes (Hayes *et al.*, 2006).

The reliability of the AAQ-ii has been shown to be good with an alpha coefficient of .84, a 3-month test-retest reliability of .81 and a 12-month test-retest reliability of .79. Studies have shown that the AAQ-ii is associated with the variables to which it is theoretically tied (Bond *et al.*, 2011).

The AAQ-ii was not designed as a tool for diagnosing mental disorders; however, Bond *et al.* (2011) were able to identify a cut-off point of 24 or above to indicate clinical levels of psychological inflexibility.

### Reliable and Clinically Significant Outcomes

Both Study 1 and 2 follow the suggestion of Delgadillo *et al.* (2014) who proposed that a comprehensive investigation of outcomes should combine effect sizes (ES) and rates of reliable and clinically significant improvement (RCSI). Reliable and clinically significant changes were used as individual outcomes for course effectiveness in this study. Reliable change refers to the extent to which change falls beyond that which would occur due to the measurement variability of a psychometric instrument (Wise, 2004). The Reliable Change Index (RCI; Jacobson & Truax, 1991) specifies the amount of change a client must show on a specific psychometric measure for that change to be reliable. The RCI for the PHQ-9 is 9, for the GAD-7 the RCI is 6, for the CORE-10 the RCI is 6 and for the AAQ-II the RCI is 9. Clinically significant change indicates whether an individual's symptoms have improved by moving from the clinical to non-clinical range (or vice versa for deterioration). The criteria for clinical change on the GAD-7 is a final score of 8 or less with pre-intervention score of 9 or more. For the PHQ-9, it is a final score of 10 or less with a pre-intervention score of 11 or more. For the CORE-10, it is a final score of 10 or below and a pre-intervention score of 11 or above. For the AAQ-II, it is a final score of 23 or below and a pre-intervention score of 24 or above. Together, the RCI and clinically significant change have been recommended as robust methods for assessing the degree of recovery following psychological interventions (Evans, Margison & Barkham, 1998; McMillan, Richards & Gilbody, 2010). It has also been proposed that rates of reliable deterioration should also be reported (Delgadillo, McMillan, Leach, Lucock, Gilbody & Wood, 2014).

Focusing only on recovery fails to recognise significant symptom changes for clients who still meet criteria for common mental health problems but who may feel considerably better. Minimal clinically important differences (MCID) are client-derived scores that reflect changes on psychometric measures that are meaningful for the client (Cook, 2008). The MCID for the GAD-7 is 4 points (Toussaint *et al.*, 2020) and for the PHQ-9 it is 5 points (Löwe, Unützer, Callahan, Perkins & Kroenke, 2004). No MCID has been reported for the CORE-10 or the AAQ-II.

Benchmarking is the statistical comparison of results found in routine clinical evaluations against those of clinical trials (high efficacy benchmarks) or those observed in control groups (no significant treatment effects; Lueger & Barkham, 2010; Minami, Serlin, Wampold, Kircher & Brown, 2008). Following the benchmarking methodology suggested by Delgadillo *et al.* (2016), ES in this study were compared to three benchmarks. For anxiety and depression scores, two benchmarks were derived from previous studies into AYL (Cartwright & Hooper, 2017; Aneurin Bevan University Health Board [ABUHB], 2019) and a third benchmark derived from a meta-analysis of guided self-help interventions for anxiety and depression (Coull & Morris, 2011). The outcomes for psychological flexibility were compared against two benchmarks from previous studies: Cartwright and Hooper (2017) and a meta-analysis of guided self-help interventions for mindfulness/acceptance (Cavanagh, Strauss, Forder & Jones, 2014). All of these benchmarks are based on group outcomes.

Another benchmark will be used to look at individual outcomes. This benchmark comes from the Improving Access to Psychological Therapies (IAPT) guidelines which state that a minimum of 50% of clients should show clinically significant change (National Collaborating Centre for Mental Health, 2020). Gyani, Shafran, Layard and Clark (2013) proposed that this number should be based on "reliable recovery" i.e., individuals who scored above the clinical cut-off at initial assessment; showed reliable improvement (based on the RCI) during treatment; and, scored below the clinical cut-offs at the end of treatment. It was necessary to use an English benchmark as no Welsh benchmark figures for individual outcomes were available.

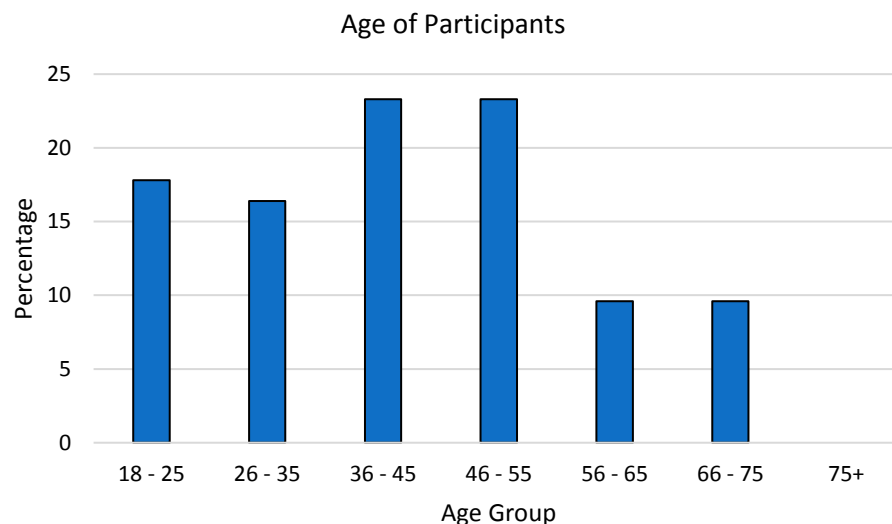
## Results

### Participant Demographics

102 participants agreed to complete outcome measures over the course of the AYL intervention. Of those that completed the demographic questionnaire, 55.4% of participants were female, 43.2% were male, 1.4% identified as non-binary and 1.4% of participants were transgender.

#### *Age of Participants*

The ages of the participants are shown in Figure 1.



*Figure 1: Age of participants who attended the AYL course*

#### *Ethnicity of Participants*

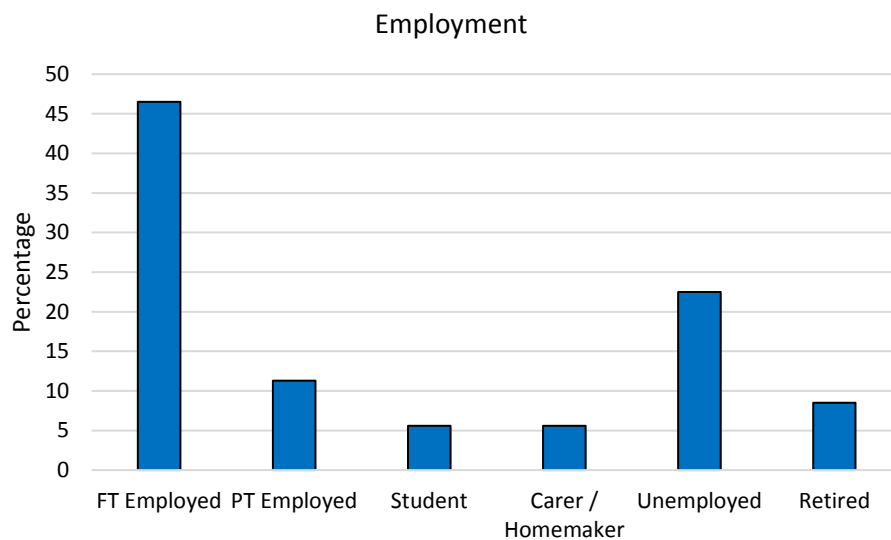
The vast majority (91.8%) of participants were White-British. 2.7% of participants identified as White-Other, 2.7% of participants stated that they were White + Black Caribbean, 1.4% of participants stated that they were Pakistani and 1.4% identified as White + Asian.

#### *Disability*

18.6% of participants described themselves as having a disability.

#### *Employment Status of Participants*

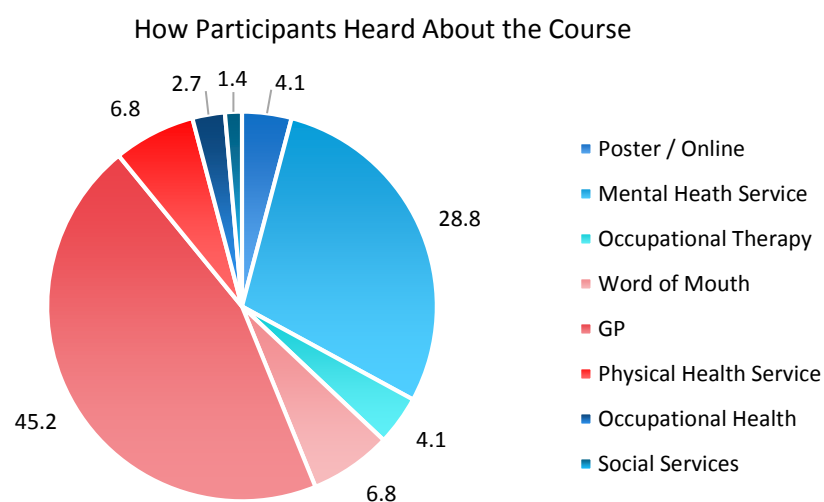
The employment status of the participants can be seen in Figure 2.



*Figure 2: Employment status of participants who attended the AYL course*

#### *How Participants Heard About the Course*

The majority of participants (45.2%) heard about the AYL course from their GP. Figure 3 shows the breakdown of all the sources that provided participants with information about the course.



*Figure 3: How participants heard about the AYL Course*

#### *Participants' Goals for the Course*

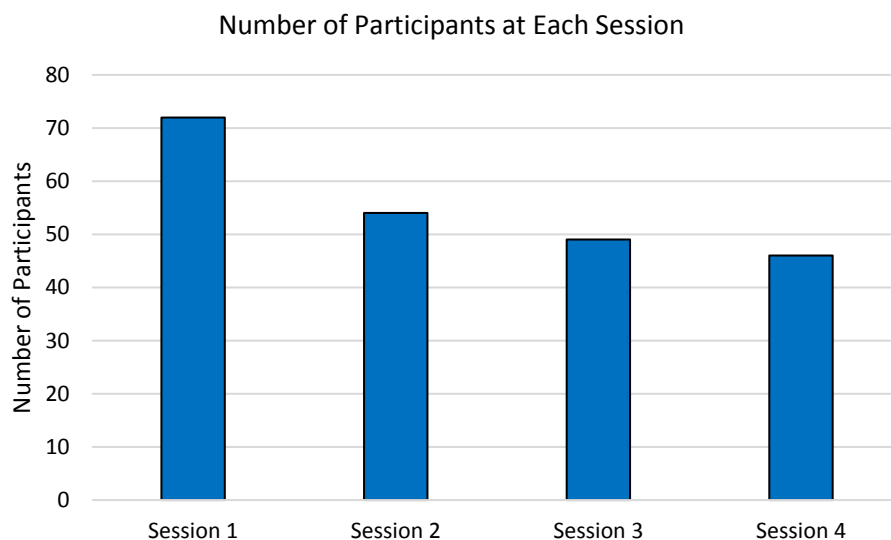
Participants were asked to identify their goals for the course session. These are shown in Figure 4.



*Figure 4: Participants' goals for the course*

#### Attendance

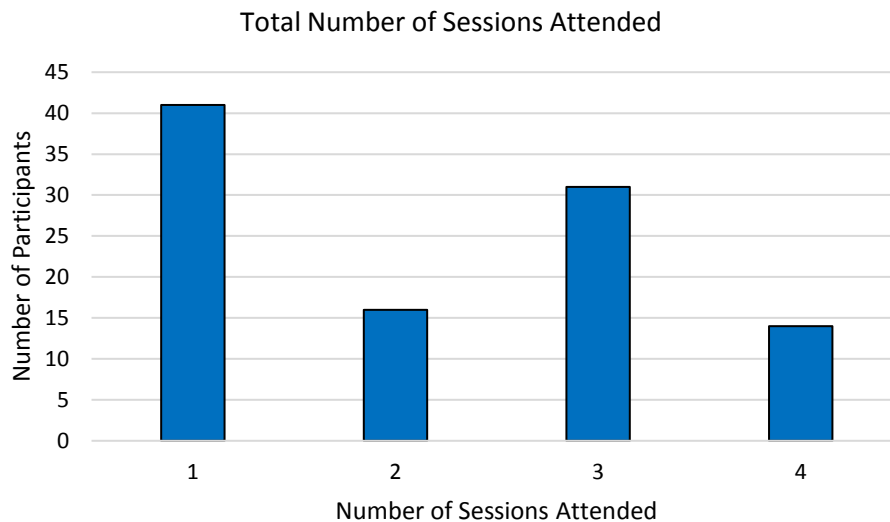
Figure 5 provides information about the number of participants in each session. It should be noted that it was not mandatory for participants to complete the questionnaires, therefore, this figure is an approximation based on the number of questionnaires completed.



*Figure 5: Number of participants in each AYL session*

Figure 6 shows the total number of sessions attended by each participant. A cut-off of attendance at three or more sessions was used to distinguish "attenders" from "drop-outs". Based on this, 43.4% of participants were considered to be attenders and 56.6% were drop-outs. 13.2% of participants attended all four sessions.





*Figure 6: Total number of sessions attended by each participant*

#### Pre-Intervention Symptoms

Table 1 shows the breakdown of pre-intervention GAD-7, PHQ-9, CORE-10 and AAQ-ii scores by severity range.

Questionnaire	Non-Clinical	Mild	Moderate	Moderate-Severe	Severe
GAD-7	6 (9.8%)	20 (32.8%)	21 (34.4%)	N/A	14 (23.0%)
PHQ-9	8 (13.3%)	11 (18.3%)	16 (26.7%)	15 (25.0%)	10 (16.7%)
CORE-10	10 (16.4%)	13 (21.3%)	12 (19.7%)	11 (18.0%)	15 (24.6%)
	<b>Non-Clinical Range</b>		<b>Clinical Range</b>		
AAQ-ii	34 (34.0%)		66 (66.0%)		

*Table 1: Breakdown of pre-intervention GAD-7, PHQ-9, CORE-10 and AAQ-ii scores by severity range*

Table 2 indicates the presentation (anxiety, depression, co-morbid) for attenders and drop-outs. Participants scoring above clinical cut-offs on both the PHQ-9 and the GAD-7 were considered to have comorbid anxiety and depression. If a participant scored above clinical cut-off on the GAD-7 and not the PHQ-9, they were considered to have an anxiety disorder (and vice versa for the PHQ-9 and depression).

There were no significant differences between attenders and drop-outs on any of the outcome measures at pre-intervention: CORE-10 ( $t(94) = 0.408$ , ns), GAD-7 ( $t(99) = 0.862$ , ns), PHQ-9 ( $t(97) = 0.399$ , ns) and AAQ-ii ( $t(99) = 0.636$ , ns).

Attendance	Presentation		Mean GAD-7 Score	Mean PHQ- 9 Score	Mean CORE-10 Score	Mean AAQ-II Score
Attendees (≥3 sessions)	Anxiety	3 (6.7%)	10.93 (SD = 5.23)	13.20 (SD = 6.40)	18.07 (SD = 8.26)	31.86 (SD = 9.61)
	Depression	3 (6.7%)				
	Co-morbid	35 (77.8%)				
	None	4 (8.9%)				
Drop-outs (≤2 sessions)	Anxiety	6 (10.5%)	11.36 (SD = 5.28)	13.22 (SD = 7.10)	19.50 (SD = 8.53)	30.79 (SD = 10.17)
	Depression	4 (7.0%)				
	Co-morbid	43 (75.4%)				
	None	3 (5.3%)				
Total	Anxiety	9 (8.9%)	11.17 (SD = 5.23)	13.21 (SD = 6.76)	18.84 (SD = 8.39)	31.26 (SD = 9.90)
	Depression	7 (6.9%)				
	Co-morbid	78 (76.5%)				
	None	7 (6.9%)				

Table 2: Pre-intervention clinical presentation of attendees and drop-outs

### Overall Group Outcomes

Table 3 shows the group outcomes for all participants. A repeated measures ANOVA was conducted to determine the impact of the number of sessions attended on symptoms levels.

The main effect of number of sessions attended was significant:  $F(1.933, 21.259) = 10.781, p < .05$ . A *post hoc* pairwise comparison using the Bonferroni correction showed that the decrease in scores reached significance when comparing the scores from session 1 with those from session 3 (17.40 vs. 12.54,  $p < .05$ ) and session 4 (17.40 vs. 11.73,  $p < .05$ ). No other pairwise comparisons were significant.

The main effect of symptom type was significant:  $F(3, 33) = 66.233, p < .001$ . The interaction between the number of sessions attended and symptom type was significant:  $F(9, 99) = 1.999, p < .05$ . Pairwise comparisons indicated that outcomes on all four questionnaires were significantly different to each other, with the most change being found on the CORE-10 and the least change being found on the GAD-7. These results are displayed in Figure 7.

	N	Pre-Interventi on Mean (SD)	Post-Interventi on Mean (SD)	Pre-Post Change Mean (SD)	95% CI range	T	Cohen's d	Effect Size
<b>Whole sample (attended 2 – 4 sessions)</b>								
PHQ-9	60	12.75 (6.38)	8.08 (6.08)	- 4.07 (5.51)	3.25 – 6.08	6.593**	0.851	Large
GAD-7	59	10.54 (5.11)	7.71 (5.30)	-2.88 (5.82)	1.32 – 4.33	3.76**	0.490	Small
CORE-10	60	18.15 (8.00)	11.82 (7.59)	- 6.50 (8.48)	4.16 – 8.51	5.832**	0.753	Moderate
AAQ-ii	60	31.47 (9.84)	26.50 (10.77)	-4.97 (7.35)	3.07 – 6.87	5.232**	0.676	Moderate
<b>Attended 3 – 4 sessions</b>								
PHQ-9	43	13.09 (6.43)	7.72 (6.39)	- 5.55 (5.77)	3.60 – 7.14	6.128**	0.934	Large
GAD-7	43	10.84 (5.33)	7.12 (5.16)	- 3.81 (5.94)	1.90 – 5.54	4.135**	0.631	Moderate
CORE-10	45	17.96 (8.20)	11.40 (8.06)	- 6.78 (9.58)	3.70 – 9.41	4.627**	0.690	Moderate
AAQ-ii	44	31.86 (9.84)	26.27 (11.12)	- 5.59 (7.65)	3.27 – 7.92	4.850**	0.731	Moderate
<b>Completed all 4 sessions</b>								
PHQ-9	14	11.36 (6.38)	6.29 (6.81)	- 5.07 (5.43)	1.94 – 8.21	3.496*	0.934	Large
GAD-7	13	5.53 (1.53)	5.77 (5.70)	- 3.25 (3.05)	1.15 – 4.85	3.541*	0.982	Large
CORE-10	14	17.07 (2.50)	9.79 (8.56)	- 4.07 (7.92)	2.99 – 11.59	3.661*	0.978	Large
AAQ-ii	14	30.36 (9.83)	23.36 (10.92)	- 7.00 (6.59)	3.20 – 10.80	3.976*	1.063	Large

\* p < 0.05, \*\* p < 0.001

Table 3: Group outcomes on all measures for attenders and drop-outs

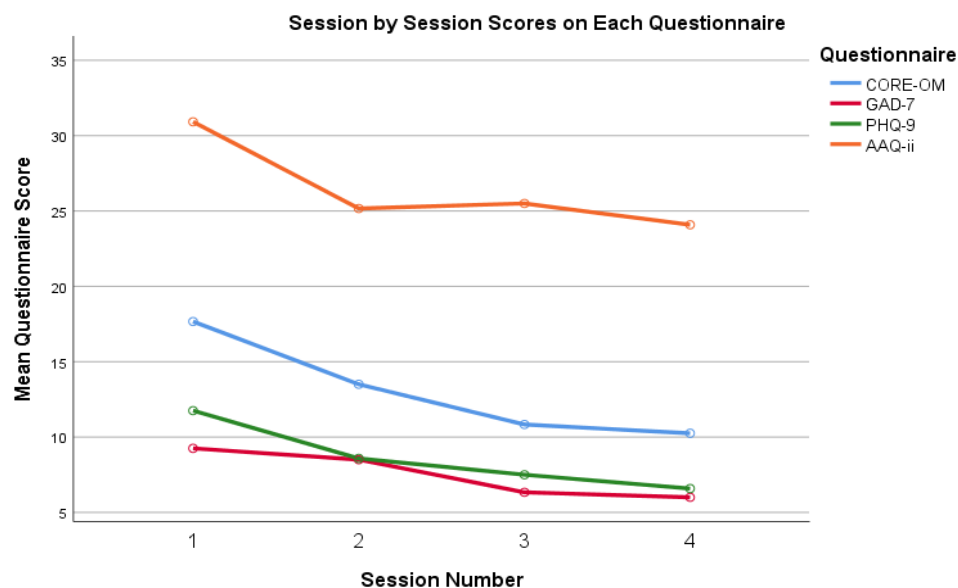
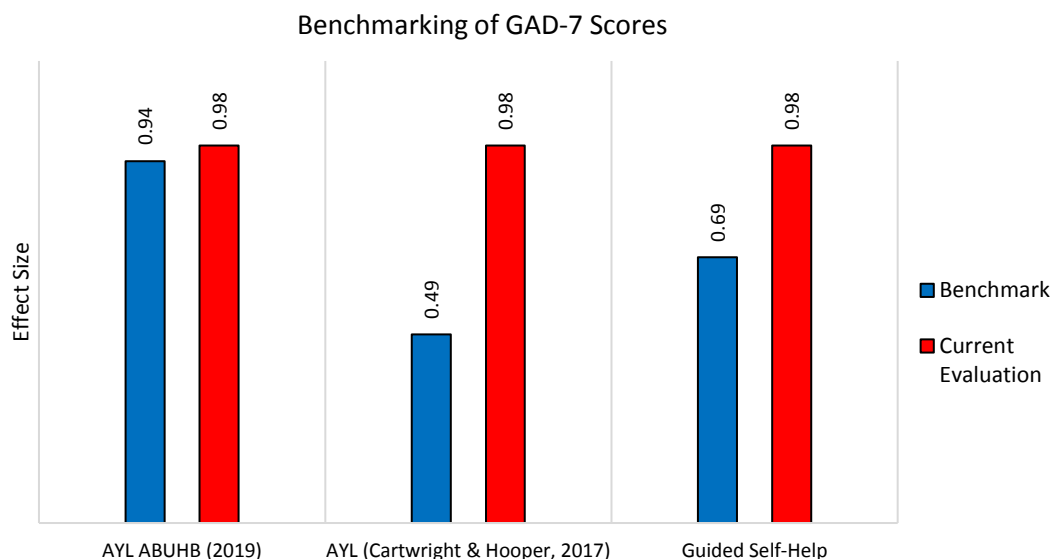


Figure 7: Session-by-session scores for each of the outcome measures

## Benchmarking

### *Anxiety (GAD-7)*

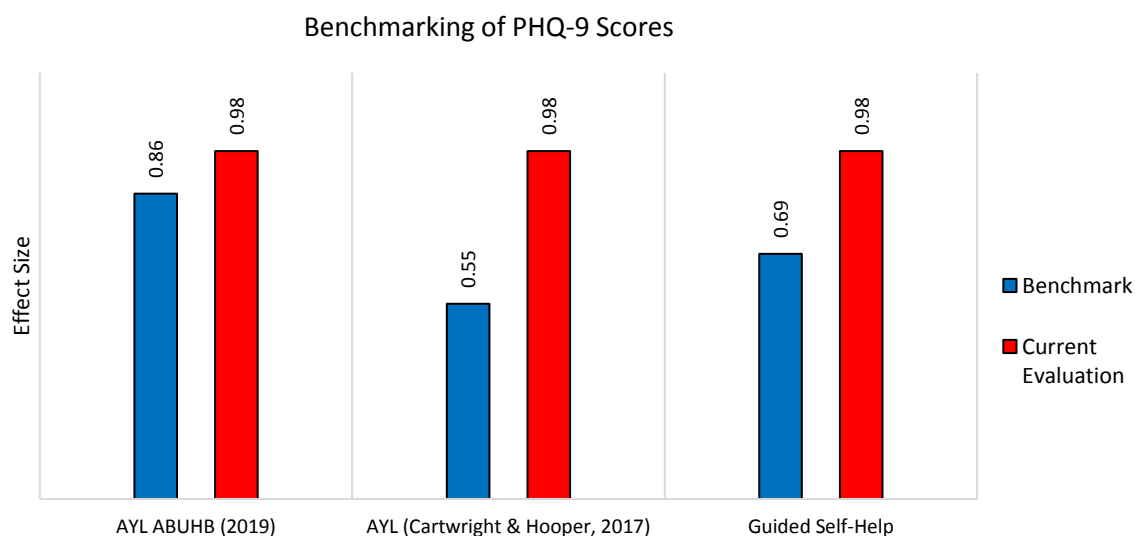
The effect size on changes in anxiety symptoms in Study 1 was similar to that reported by ABUHB (2019) and significantly larger than the effect sizes reported in the other two benchmark studies (Cartwright & Hooper, 2017; Coull & Morris, 2011). These results are shown in Figure 8.



*Figure 8: GAD-7 effect sizes for participants who attended all four sessions compared to benchmarks*

### *Depression (PHQ-9)*

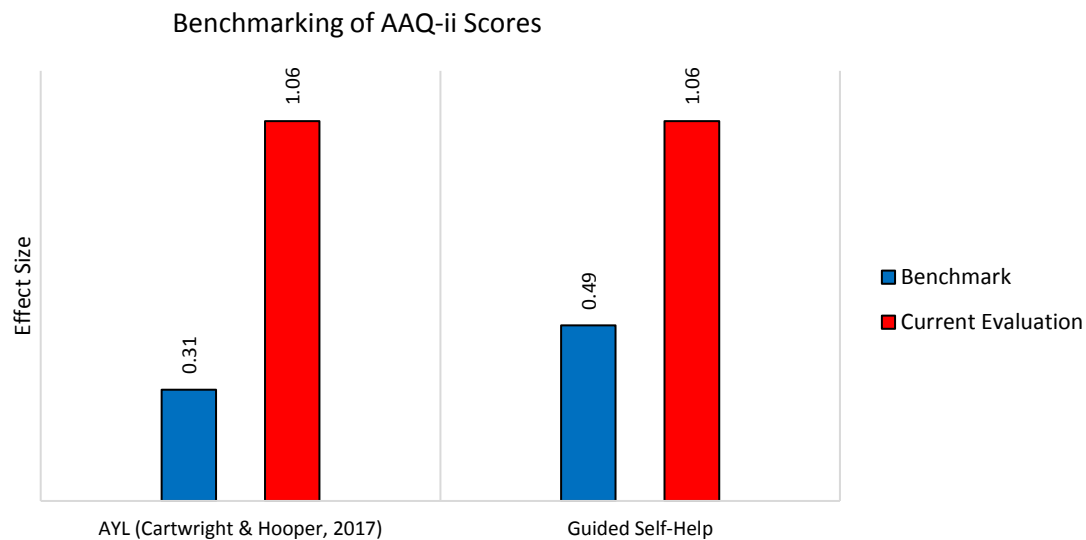
The results from Study 1 showed a slightly larger effect size for changes in depression symptoms than that reported by ABUHB (2019) and a significantly larger effect size on changes in depression symptoms than reported in the other two benchmark studies (Cartwright & Hooper, 2017; Coull & Morris, 2011). These results are shown in Figure 9.



*Figure 8: PHQ-9 effect sizes for participants who attended all four sessions compared to benchmarks*

### *Psychological Flexibility (AAQ-II)*

The benchmark comparisons for psychological flexibility indicated that Study 1 showed a significantly larger effect size on changes in psychological flexibility than the results reported by both Cartwright and Hooper (2017) and Cavanagh *et al.* (2014). These results are shown in Figure 10.



*Figure 10: AAQ-II effect sizes for participants who attended all four sessions compared to benchmarks*

### Overall Individual Outcomes

Individual outcomes were calculated using a number of measures: clinically reliable change, clinically significant change and MCID. These results can be seen in Table 4.

	N	Positive Clinically Significant Change	Positive MCID	Reliable Recovery	Stasis	Negative Clinically Significant Change	Reliable Deterioration	Negative MCID
Whole sample (attended 2 – 4 sessions)								
PHQ-9	60	12 (20.0%)	12 (20.0%)	29 (48.3%)	49 (81.7%)	0 (0.0%)	0 (0.0%)	2 (3.3%)
GAD-7	58	21 (36.2%)	27 (46.6%)	17 (29.3%)	36 (62.1%)	3 (5.2%)	3 (5.2%)	8 (13.8%)
CORE-10	60	30 (50.0%)	N/A	17 (28.3%)	27 (45.0%)	3 (5.0%)	3 (5.0%)	N/A
AAQ-II	60	16 (26.7%)	N/A	8 (13.3%)	31 (51.7%)	2 (3.3%)	1 (1.7%)	N/A
Attended 3 – 4 sessions								
PHQ-9	44	10 (22.7%)	27 (61.2%)	10 (22.7%)	34 (77.3%)	0 (0.0%)	0 (0.0%)	2 (4.5%)
GAD-7	41	20 (48.9%)	22 (53.7%)	17 (41.5%)	19 (46.3%)	2 (4.9%)	2 (4.9%)	5 (12.2%)
CORE-10	45	21 (46.7%)	N/A	17 (37.8%)	21 (46.7%)	3 (6.7%)	3 (6.7%)	N/A
AAQ-II	44	12 (27.3%)	N/A	6 (13.6%)	31 (70.5%)	1 (2.3%)	1 (2.3%)	N/A
Completed all 4 sessions								
PHQ-9	14	2 (14.3%)	9 (64.3%)	2 (14.3%)	9 (64.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
GAD-7	13	4 (30.8%)	6 (46.2%)	3 (23.1%)	7 (53.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
CORE-10	14	5 (35.7%)	N/A	4 (28.6%)	9 (64.3%)	0 (0.0%)	0 (0.0%)	N/A
AAQ-II	14	5 (35.7%)	N/A	3 (21.4%)	9 (64.3%)	0 (0.0%)	0 (0.0%)	N/A

Table 4: Individual outcomes for all participants based on number of sessions attended

### Benchmarking

The benchmark provided by the English IAPT services is that 50% of clients should show reliable recovery. For the current evaluation, this figure is below 50% for all outcome measures regardless of the number of sessions attended by the participants (see Table 4).

### The Impact of Initial Symptom Severity on Outcome

#### *Group Outcomes by Initial Symptom Severity*

Pre- and post-intervention scores were analysed to determine whether initial symptom severity had an impact on the effectiveness of the AYL intervention for anxiety, depression, psychological flexibility and levels of general psychological distress. This information is displayed in Table 5.

The results of independent t-tests on GAD-7 scores indicated that individuals who scored as moderately-severe to severely anxious prior to intervention showed a significantly greater reduction in anxiety than those categorised with mild to moderate anxiety ( $t(56) = 3.556, p < .05$ ).

The results of independent t-tests on PHQ-9 scores showed that individuals who scored as moderately-severe to severely depressed prior to the intervention showed a significantly greater reduction in depression than those categorised with mild to moderate anxiety ( $t(50) = 2.602, p < .05$ ).

The results of independent t-tests on CORE-10 scores showed that individuals whose symptoms were moderately-severe to severe prior to the intervention showed a significantly greater reduction in symptoms than those with mild to moderate symptoms ( $t(58) = 4.556, p < .001$ ).

The results of independent t-tests on AAQ-II scores showed that individuals whose degree of psychological inflexibility was in the clinical range prior to the intervention showed a significantly greater reduction in symptoms than those in the non-clinical range ( $t(58) = 2.807, p < .05$ ).

Severity	N	Pre-Intervention Mean (SD)	Post-Intervention Mean (SD)	Pre-Post Change Mean (SD)	95% CI range	t	Cohen's d	Effect Size
<b>GAD-7</b>								
Mild - Moderate	38	7.82 (2.87)	6.74 (4.78)	1.08 (5.06)	- 0.59 – 2.74	1.313 (n.s.)	0.21	Small
Moderately Severe - Severe	20	16.25 (2.94)	9.95 (5.56)	-6.30 (5.73)	3.62 – 8.98	4.916**	1.10	Large
<b>PHQ-9</b>								
Mild - Moderate	25	7.91 (3.77)	5.27 (4.17)	-2.70 (4.00)	1.25 – 4.02	3.880**	0.68	Moderate
Moderately Severe - Severe	27	18.78 (2.51)	14.22 (4.82)	-4.56 (4.03)	2.55 – 6.56	6.039**	1.16	Large
<b>CORE-10</b>								
Mild - Moderate	34	12.41 (4.99)	9.82 (5.59)	-2.59 (6.85)	0.20 – 4.98	2.205*	0.38	Small
Moderately Severe - Severe	26	25.65 (3.86)	14.42 (9.06)	- 11.23 (7.82)	8.07 – 14.39	7.323**	1.44	Large
<b>AAQ-II</b>								
Non-Clinical	20	19.90 (5.40)	18.50 (8.93)	-1.40 (5.79)	- 1.30 – 4.10	1.807 (n.s.)	0.24	Small
Clinical	40	37.25 (5.41)	30.50 (9.35)	-6.75 (7.47)	4.36 – 9.14	5.713**	0.90	Large

\*  $p < 0.05$  \*\*  $p < 0.001$

Table 5: Post-intervention group outcomes by initial symptom severity

#### Individual Outcomes by Initial Symptom Severity

A variety of individual outcome measures were used to determine whether initial symptom severity had an impact in the effectiveness of the AYL intervention for anxiety; depression; levels of general distress; and, psychological flexibility. This information is displayed in Table 6.

	N	Positive Clinically Significant Change	Positive MCID	Reliable Recovery	Stasis	Reliable Deterioration	Negative Clinically Significant Change	Negative MCID
<b>GAD-7</b>								
Mild – Moderate	38	9 (23.7%)	12 (31.6%)	9 (23.7%)	25 (65.8%)	4 (10.5%)	4 (10.5%)	7 (18.4%)
Moderately Severe – Severe	20	12 (60.0%)	15 (75.0%)	8 (40.0%)	8 (40.0%)	0 (0.0%)	0 (0/0%)	15 (75.0%)
<b>PHQ-9</b>								
Mild - Moderate	25	3 (12.0%)	10 (40.0%)	2 (8.0%)	22 (88.0%)	0 (0.0%)	0 (0.0%)	1 (4/0%)
Moderately Severe - Severe	27	9 (33.3%)	19 (70.4%)	9 (33.3%)	18 (66.7%)	0 (0/0%)	0 (0/0%)	1 (3.7%)
<b>CORE-10</b>								
Mild - Moderate	34	9 (26.5%)	N/A	9 (26.5%)	22 (64.7%)	3 (8.8%)	3 (8.8%)	N/A
Moderately Severe - Severe	26	21 (80.8%)	N/A	11 (42.3%)	5 (19.2%)	0 (0.0%)	0 (0.0%)	N/A
<b>AAQ-ii</b>								
Non-Clinical	20	1 (5.0%)	N/A	0 (0.0%)	17 (85.0%)	2 (10.0%)	1 (5.0%)	N/A
Clinical	40	15 (37.5%)	N/A	9 (22.5%)	25 (62.5%)	0 (0.0%)	0 (0.0%)	N/A

Table 6: Post-intervention individual outcomes by initial symptom severity

Chi-square tests were used to determine whether reliable recovery rates were impacted by the participants' initial symptom severity.

Differences between reliable recovery rates for individuals in the moderately-severe to severe anxiety category compared to those in the mild to moderate anxiety category were not statistically significant ( $\chi^2(1) = 0.273$ ,  $p = .602$ ). See Figure 11.

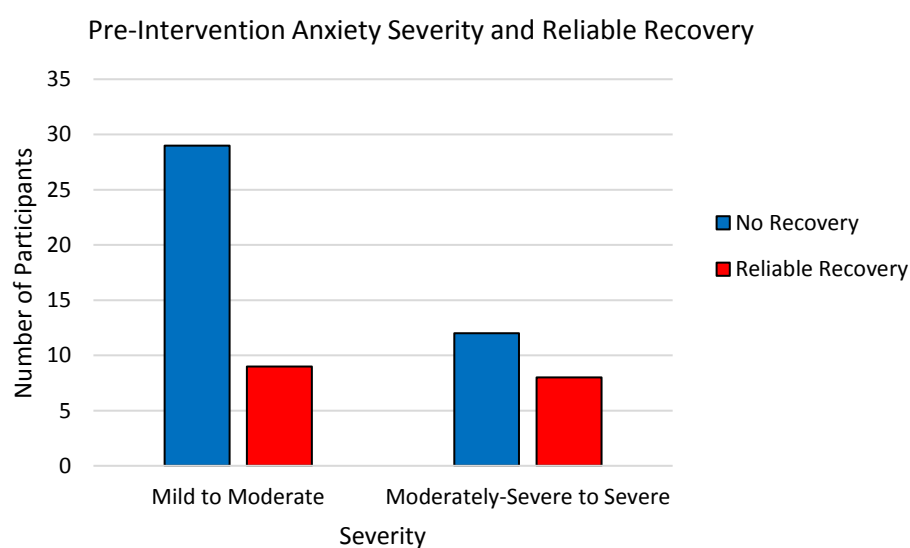


Figure 11: Number of participants who demonstrated reliable recovery on the GAD-7 calculated by initial symptom severity



Differences between reliable recovery rates for individuals with moderately-severe to severe depression were not significantly different to those for individuals in the mild to moderate depression category ( $\chi^2(1) = 3.328$ ,  $p = .068$ ). See Figure 12.

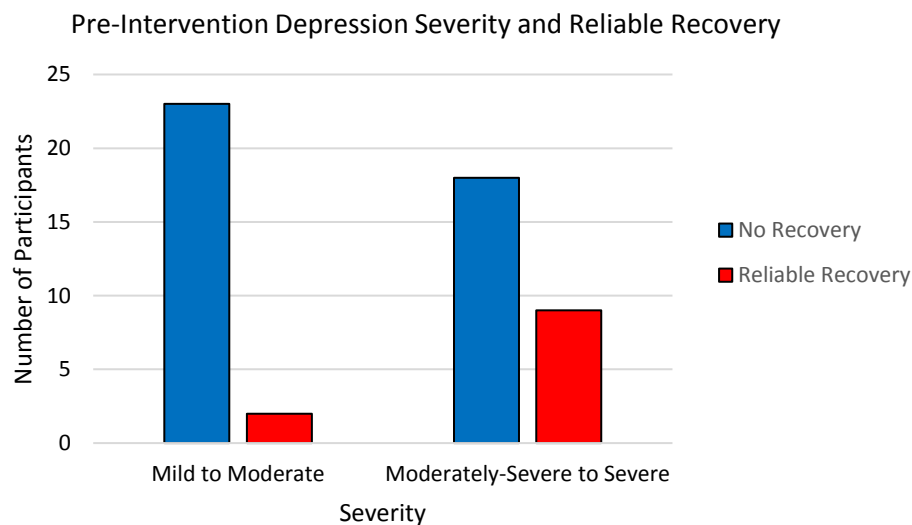


Figure 12: Number of participants who demonstrated reliable recovery on the PHQ-9 calculated by initial symptom severity

Reliable recovery rates for general psychological distress between individuals in the moderately-severe to severe category and those in the mild to moderate category were not statistically significant ( $\chi^2(1) = 1.663$ ,  $p = .197$ ). See Figure 13.

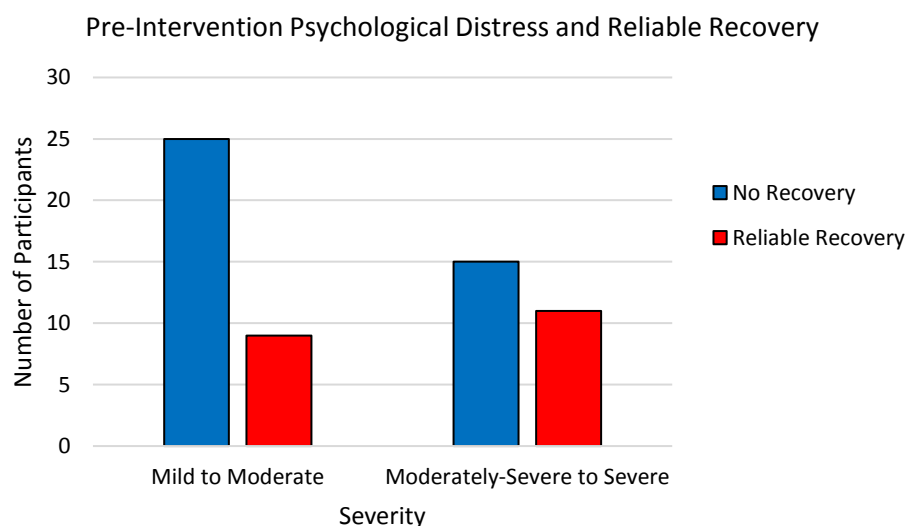
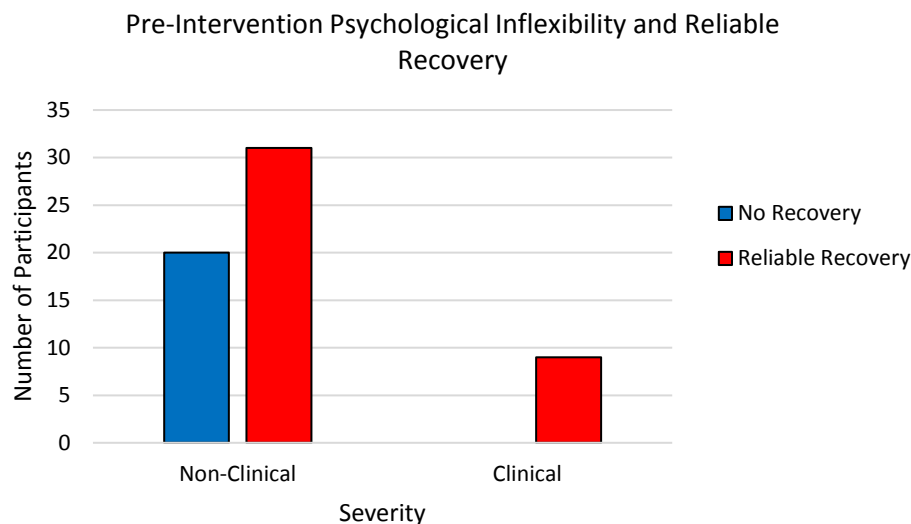


Figure 13: Number of participants who demonstrated reliable recovery on the CORE-10 calculated by initial symptom severity

Reliable recovery rates for psychological inflexibility between individuals in the clinical category and those in the non-clinical category were statistically significant ( $\chi^2(1) = 4.615$ ,  $p = <.05$ ). See Figure 14.

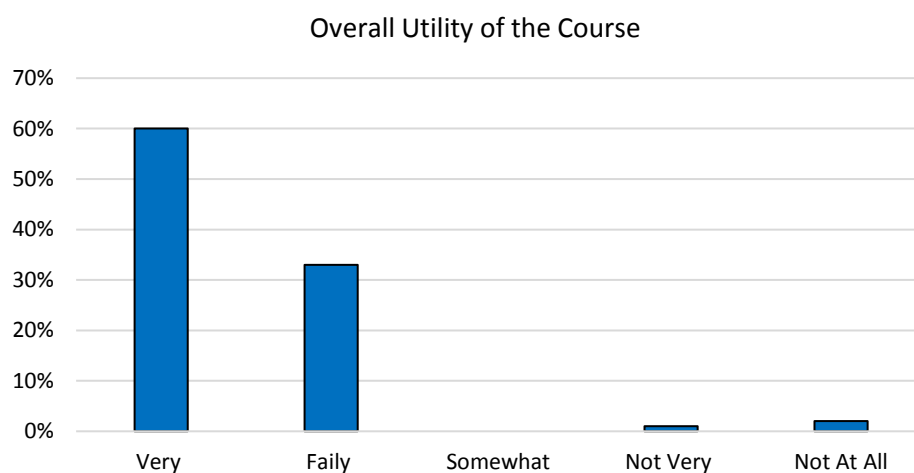


*Figure 14: Number of participants who demonstrated reliable recovery on the AAQ-II calculated by initial symptom severity*

#### Feedback from Participants

95.6% of attendees who responded stated that they would recommend the AYL course to others.

Participants were asked to indicate how useful they found the course overall. These results can be seen in Figure 15.



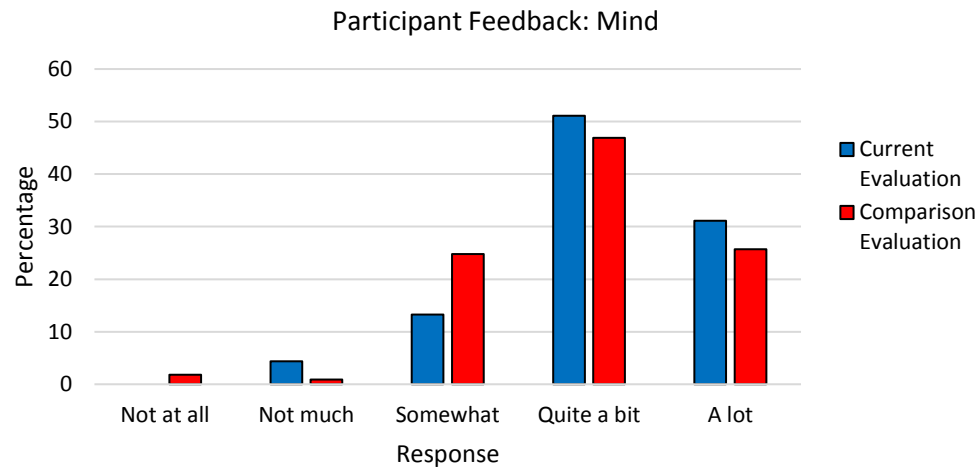
*Figure 15: Participants' feedback regarding the utility of the course*

Participants were asked to rate how well they felt the course covered the main topics that AYL is designed to include on a 5-point Likert scale (from "Not at all" to "A lot"). The results are shown in Table 7.

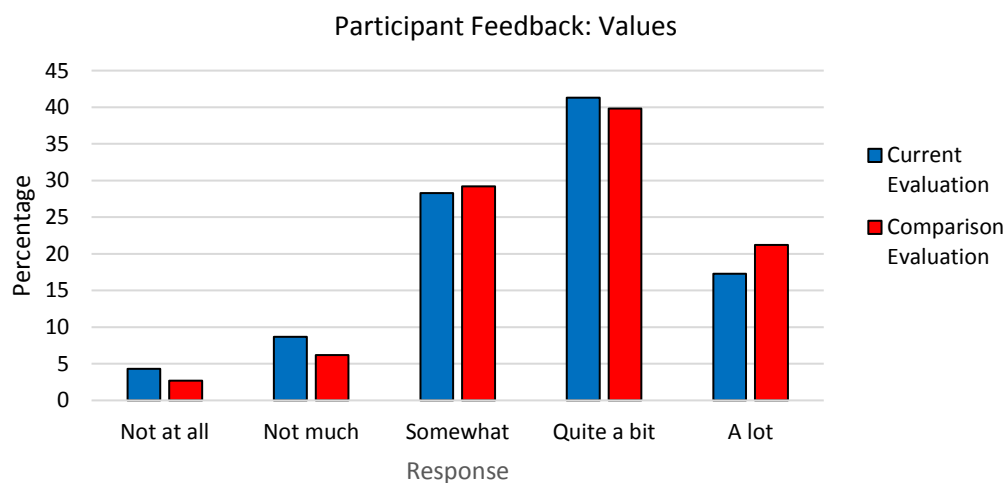
Course Focus	Not at all	Not much	Somewhat	Quite a bit	A lot
Mind	0.0%	4.4%	13.3%	51.1%	31.1%
Values	4.3%	8.7%	28.3%	41.3%	17.4%
Thoughts	0.0%	4.3%	17.4%	45.7%	32.6%
Actions	0.0%	8.7%	21.7%	47.8%	21.7%

*Table 7: Participants' feedback regarding the coverage of the main aspects of the AYL course*

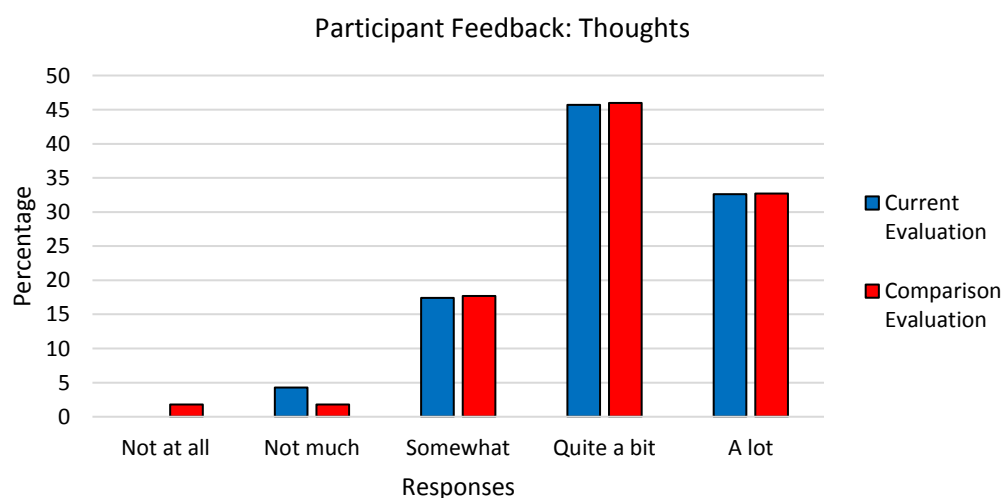
This feedback was compared to the feedback obtained from participants in the ABUHB (2019) AYL course evaluation as a form of benchmarking. The results can be seen in Figures 16 – 19. It can be seen from these graphs that the outcomes for Study 1 were similar to, or better, than those reported by ABUHB.



*Figure 16: Comparison between Study 1 and ABUHB (2019) for understanding of mind*



*Figure 17: Comparison between Study 1 and ABUHB (2019) for understanding of values*



*Figure 18: Comparison between Study 1 and ABUHB (2019) for understanding of thoughts*

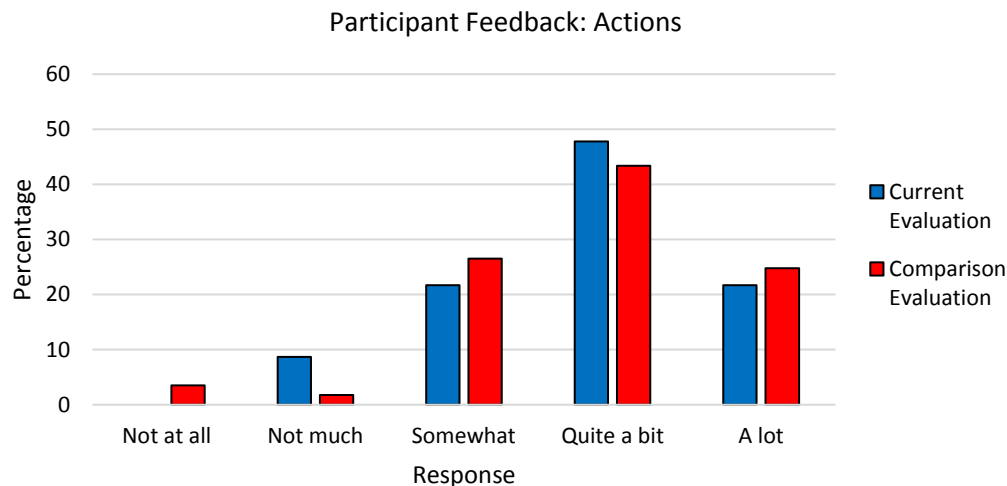


Figure 19: Comparison between Study 1 and ABUHB (2019) for understanding of actions

The positive themes that came from participant feedback on the Evaluation Questionnaire were:

- Mindfulness:
  - *"Understanding I don't have to follow my mind"*
  - *"I have learned to focus more on certain things and to know not to listen to the mind."*
  - *"The basic assertion that you are not your mind and that we are free to ignore our thoughts if we choose to"*
  - *"The descriptions of different mindfulness activities - getting to try them out"*
  - *"The concept that you are not your mind"*
- Expressed gratitude for the course and the facilitators:
  - *"Although I could read the screen it was also helpful to have people read it as well. It [the information] seemed to sink in"*
  - *"While a lot of the information wasn't unfamiliar it's been a really useful way of looking at things"*
  - *"Very informal and initially that's what I needed in order to take things in. It's saved my life"*
  - *"Welcoming and easy going"*
  - *"Interesting concepts"*
- Specific exercises and techniques:
  - *"I found the defusion techniques very helpful"*
  - *"Time outs, spending time focusing on yourself, surrounding noises, leaves on stream helpful to me when my brain refuses to switch off"*
  - *"Guidance on developing a healthy scepticism for particular thoughts"*
- The usefulness of the handouts/homework activities.
  - *"Things to do at home have been very useful"*
  - *"Handouts and homework tasks [are useful]"*
  - *"The exercises to use at home when stressed"*
- Developing a different understanding / perspective:
  - *"Knowing how the mind works"*
  - *"Learning about thoughts, rumination"*
  - *"[Understanding] how the mind makes you change your actions"*

- *"Explaining how and why the mind works as it does - common thought processes and evolutionary context"*
- *"How to think of things with a different perspective and let go of things out of my control"*
- A general positive impact of attending the course on quality of life:
  - *"I have slept much better since using this"*
  - *"Looking on the brighter side"*
  - *"The power of acceptance"*
  - *"[It is useful to know that] other people have unhelpful thoughts just like me"*
  - *"Helping to deal with negative thoughts"*

It should be noted that very few participants (n = 12) reported any negative themes. Those that were reported fell into two categories:

- Method of presentation:
  - *"Presentation for 2 hours. Although very useful would be good if some activities"*
  - *"Simply reading the slides...uninspiring"*
  - *"The lack of interaction, the poor quality Powerpoint, I found the examples used impossible to relate to"*
  - *"Summary at the end of the session because it was in the handout and start of next session"*
  - *"Auto suggestive negative statements being recited and shown on screen e.g. 'I'm worthless', 'I don't fit in' etc. Although contextually these are given as examples of thought processes, I'd say they are unnecessary and serve to actually plant those thoughts in the minds of attendees"*
  - *"The format of slides and reading aloud - no camaraderie and little active learning"*
- Specific techniques:
  - *"Closing our eyes and losing ourselves in thought (leaf on the water) etc. not for me"*
  - *"Anger and defusion"*

#### Feedback from Facilitators

The facilitators were positive about the idea of delivering psychoeducational courses which they felt were beneficial in reducing the LPMHSS waiting list and referrals to other mental health services. Staff were also enthusiastic about AYL as they felt that the course material was better at keeping people engaged than Stress Control, the other psychoeducational course run by LPMHSS. It was suggested that this may have been due to the practical mindfulness exercises that were taught. The facilitators felt that the Stress Control course was better suited for individuals who were finding it difficult to cope, whereas AYL was more effective for individuals who had difficulties in life that they could not change and those who had high levels of rumination.

It was recognised that, due to staff turn-over, no one currently presenting the AYL course had been specifically trained in AYL. This resulted in the facilitators reading the slides off the screen without feeling confident to elaborate on the ideas presented. New members of staff were asked to observe the AYL course before being required to deliver the course materials themselves.

The facilitators highlighted the fact that the handouts for the course were not available in Welsh.

The facilitators expressed concern about the limitations of the potential venues for psychoeducational courses because it was necessary to find venues that were free or available for a nominal fee. They

explained that this had an impact on the size of the groups that could be invited to the sessions. The limitations in venues also presented problems in terms of accessibility with parking being limited for some settings. Staff also felt that holding the courses on general or psychiatric hospital sites created a barrier to attendance for some participants because they felt that their situation was being “*medicalised*” and also made them more concerned about stigma related to being viewed as having a mental health problem.

LPMHSS staff also highlighted the fact that there were difficulties recruiting participants for both the Stress Control and the AYL courses because the majority of people were told by the GP that they would be offered a one-to-one intervention from LPMHSS. This led to people declining a group intervention or the GPs stating on the referral form that the individual did not want to participate in a group intervention. Given that the majority of people are referred by their GP surgery, it was suggested that it may be useful to “*rebrand*” LPMHSS with GPs, so that they are aware that psychoeducational groups will often be the first intervention offered and that one-to-one interventions are not necessarily offered are standard.

## **Discussion**

### Summary of Results

The overall group results indicated that there was a statistically significant impact of the intervention on all measures. Furthermore, the current AYL course run in HDUHB is equally or more effective than other studies against which the results were benchmarked.

The group outcomes indicated that improvements in symptoms were significantly more likely to occur for those individuals who attended three or more sessions than for those who attended less than three sessions. People who attended all four sessions showed a large effect size with respect to improvement on all four outcome measures (depression, anxiety, general psychological distress and psychological flexibility). In contrast, the outcomes for those who attended fewer sessions showed a large effect size for improvements in depression symptoms; a small effect size for improvements in anxiety symptoms; and, a moderate effect size for improvements in general psychological distress and psychological flexibility.

When looking at the data on an individual level, the majority of participants did not show any reliable recovery on any of the outcome measures used, with reliable recovery rates for individuals who attended all four sessions being between 8% and 42%. This is below the benchmark figure of 50% from IAPT England. Despite this, the MCID scores indicate that between 61% and 64% of individuals who attended three or more sessions experienced improvements in depression symptoms that were likely to have a positive impact on their quality of life. For anxiety symptoms, this figure was between 46% and 53%. Only a small number of participants showed any clinical deterioration, suggesting that the intervention did not cause harm to the participants.

Both the group and individual outcome data indicated that participants with more severe initial symptoms showed significantly greater improvement than those with milder initial symptoms. For group outcome measures, individuals with moderately-severe to severe initial symptoms showed a statistically significant ( $p < .01$ ) improvement with a large effect size for all outcome measures. Individuals with mild to moderate initial symptoms, showed no significant improvement for anxiety or psychological flexibility and a smaller effect size than seen in those with more severe initial symptoms. In terms of the individual data, this difference reached statistical significance for the CORE-10 and the PHQ-9 but not for the other two outcome measures. It is possible that this result is due to floor effects for those in the latter group i.e. individuals with mild or sub-clinical symptoms are less likely to show a symptoms reduction above the RCI of the measure.

The attrition rate of participants in the study was high, with only 13% of participants attending all four sessions of the course, although it should be noted that this is only an estimate as the figures are based on the number of completed evaluation forms. It is likely that some individuals attended all four sessions but chose not to complete the forms. Attrition rates and missed sessions for group interventions are particularly important because sessions that are missed are essentially lost forever. In contrast, individual interventions can be more flexible in the pace at which material is presented (Kellet, Clarke & Matthews, 2007).

The qualitative feedback from participants indicated that 95% would recommend the AYL course to others. Furthermore, 60% of participants found the course *“Very Useful”* with a further 33% finding it *“Fairly Useful”*. Participants felt that the course provided good information regarding the *“Mind”* and *“Thoughts”* sections of the course but that the information was not as clear for the *“Values”* and *“Actions”* sections. This suggests that the course provides a good grounding in some aspects of ACT such as mindfulness and acceptance, but that it is less clear on other important ACT principles such as identifying, and living by, personal values.

The data collected from the open-ended questions on the Evaluation Questionnaire highlighted the positive themes of mindfulness; expressed gratitude for the course and the facilitators; specific ACT techniques; a positive impact on quality of life; and, the usefulness of the materials that were given to participants. The negative themes that were identified were the method of presentation and specific ACT techniques.

The qualitative feedback from facilitators indicated that they were positive about offering psychoeducational courses in general but that they had concerns about the fact that current facilitators have not been formally trained in the delivery of AYL. This resulted in facilitators simply reading from the slides, something that participants reported as being detrimental to their engagement in the course. Concerns were also expressed about the difficulties in finding suitable cheap or free venues for both psychoeducational courses and the way in which GPs understood the interventions that LPMHSS offered.

#### Limitations of the Evaluation

Participants were not randomly selected and so caution in the interpretation of the results is necessary. Furthermore, the Evaluation Questionnaire was only completed by those individuals who attended the final session who, on the whole, were individuals considered to be attenders. This may have introduced a degree of bias into the feedback obtained and provides little indication of why individuals dropped out of the course.

The design of this study did not provide an opportunity for follow-up data to be collected so it is not possible to know whether any gains achieved were maintained over the medium to long-term.

The pre-post treatment effect sizes described in this study offer a general estimate of the *“real world”* effectiveness of AYL interventions delivered in HDUHB. Results were not analysed relative to a control group, therefore, it is possible that regression to the mean (i.e. natural fluctuations in mental health symptoms due to the passage of time) may have partly accounted for some of the reported effects. Furthermore, the data do not disentangle specific AYL treatment effects from effects that may be due to general contact with healthcare practitioners and other participants in a group-based setting.

## Conclusion

The AYL course in HDUHB performed well in comparison to benchmark studies and when looking at the group outcomes. In terms of individual outcomes, there was a lack of clinically significant positive outcomes for the vast majority of participants. However, based on MCID scores, the majority of people who attended three or more sessions achieved an improvement in depression and anxiety symptoms that were likely to have a positive impact on their quality of life. Outcomes are likely to be improved by training staff to facilitate AYL and utilising the newest version of the course. Providing staff with up-to-date training is vital because it has been shown that the characteristics and training of group facilitators determine their capacity to lead a group (Borek & Abraham, 2018). Furthermore, it has been suggested that variability in facilitators' competence may partly explain differences in outcomes of group interventions (Burlingame, Strauss & Joyce, 2013), emphasising the need for group facilitators to be offered regular training.

This evaluation showed that participants had high levels of satisfaction with the course; however, attrition rates were high. Analysis of the impact of attendance showed that outcomes were statistically more positive when clients attended three or more sessions. This illustrates the importance of supporting participants to attend as many of the sessions as possible.

It was found that the outcomes measured by the four psychometric questionnaires were significantly different. The largest improvement was measured on the CORE-10. The CORE-10 provides a measure of psychological distress associated with mental health difficulties; whereas the GAD-7 and PHQ-9 measure changes in symptoms. This may indicate that attendance at the AYL course, as delivered in HDUHB, reduced the amount of distress that participants felt about their symptoms more than it reduced the symptoms themselves. This hypothesis is supported by the relatively high MCID scores that were reported.



## **STUDY TWO: COMPARISON OF CURRENT EVALUATION DATA WITH PAST EVALUATIONS WITHIN HDUHB**

### **Introduction**

LPMHSS practitioners in Ceredigion collected pre- and post-intervention data for the AYL course run in the county for a number of years; however, this data had not been analysed. Study 2 involved analysing this data to determine the effectiveness of the AYL course. Furthermore, where possible, data from Study 2 was compared to that obtained in Study 1 in order to identify whether there have been any significant changes in the effectiveness of the AYL course over time.

### **Method**

#### Intervention

The intervention is described in Study 1 (see p. 3).

#### Design

A mixed methods design was employed. The strategy involved primary quantitative methods and secondary qualitative methods used to obtain the views of the course participants. Pre- and post-measures were used for the quantitative evaluation. Dependent t-tests on the pre- and post-intervention changes on the psychometric measures were used to calculate the effectiveness of the intervention. Calculations of clinical significance and reliable recovery rates were carried out and descriptive statistics were produced for the demographic data.

#### Questionnaires

##### *Initial and Evaluation Questionnaires*

See page 4 for description. It should be noted that this Evaluation Questionnaire did not provide open-ended questions for participants to complete.

##### *GAD-7 (Spitzer et al., 2006)*

See page 4 for description.

##### *PHQ-8 (Kroenke, Strine, Spitzer, Williams, Berry & Mokdad, 2008)*

The PHQ-8 consists of eight items each on a four-point Likert scale. The PHQ-8 includes the same questions as the PHQ-9 (see page 4 for description) without the question about risk. Studies have indicated that the PHQ-8 has Cronbach's  $\alpha$  of 0.88 and that there is no statistically significant difference between the PHQ-8 and PHQ-9 (e.g., Shin, Lee, Han, Yoon & Han, 2019).

##### *The Acceptance and Action Questionnaire (Hayes et al., 2006)*

See page 5 for description. It should be noted that the questionnaire used to collect data in this study was the original AAQ (Hayes et al., 2006). Some of the questions from the AAQ were removed to form the AAQ-ii (Bond et al., 2011). In this study, AAQ scores were converted to AAQ-ii scores for this study in order to make clear comparisons between the results obtained in Study 1 and Study 2.

*Mindfulness-Based Self-Efficacy Scale - Revised (MSES-R; Cayoun, Francis, Kasselis & Skilbeck, 2012)*

This 22-item self-report measure is designed to assess the changes in levels of perceived self-efficacy before, during and after mindfulness-based therapy. There are six subscales: Emotion Regulation, Distress Tolerance, Equanimity, Taking Responsibility, Social Skills and Interpersonal Effectiveness. Items are rated on a 5-point Likert scale. Higher scores indicate higher mindfulness-based self-efficacy. Subscales can be used separately to assist in identifying relative strengths and weaknesses or combined to provide a global score. Study 2 used only the first three subscales of the MSES-R (Emotion Regulation, Distress Tolerance and Equanimity) along with a total score calculated from these three subscales. The MSES-R total shows high internal consistency (Francis & Cayoun, 2011).

*Rosenberg Self-Esteem Scale (RSES, Rosenberg, 1965)*

The RSES is one of the most widely used measures of global self-esteem (SE; Byrne, 1996). SE has been defined as one's overall sense of worth as an individual (Rosenberg, 1965). SE has been shown to play an important role in the experience of anxiety, other affective states and interpersonal style (Greenberg *et al.*, 1992; Lightsey, Burke, Ervin, Henderson & Yee, 2006; Neustadt, Chamorro-Premuzic & Furnham, 2006; Torrey, Mueser, McHugo & Drake, 2000).

The RSES is a brief, self-completion questionnaire with ten questions each answered using a four-point Likert scale. The RSES has been shown to have test-retest correlations typically in the range of .82 to .88, and Cronbach's alpha of .88 (Rosenberg, 1965). Test-retest reliability at 2-week interval was calculated at 0.85 (Silber & Tippet, 1965). The RCI for the RSES is 3.

Total scores between 15 and 25 are considered to be in the normal range, with scores of 14 or below representing low self-esteem and scores of 26 or above being considered excessively high self-esteem.

Benchmarking

The same benchmarks were used in this study as in Study 1. In addition, the results of this study were also benchmarked against the result in Study 1.

**Results**

Participant Demographics

221 participants agreed to complete outcome measures over the course of this study. Of these, 68.1% were female and 31.9% were male. In comparison, Study 1 had a more equal representation of male and female participants (55.4% female, 43.2% male). The Demographics Questionnaire in Study 2 did not provide options for participants to identify as non-binary or transgender, therefore this information cannot be compared.

The current and previous evaluations used different age ranges. Figure 20 shows the age range of participants in Study 2.

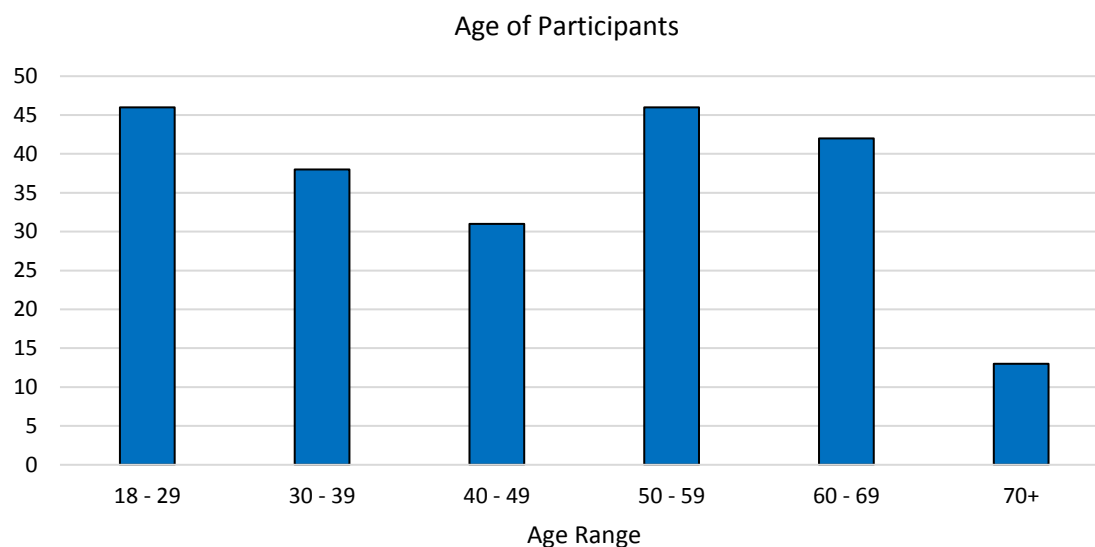


Figure 20: Age of participants

### Ethnicity of Participants

In both Study 1 and Study 2, the vast majority of participants were White-British (91.8% in Study 1 and 95.8% in Study 2). In Study 2, 4.2% of participants identified as White-Other compared to 2.7% in Study 1. In Study 1 participants also identified as White + Black Caribbean (2.7%), Pakistani (1.4%) and White + Asian (1.4%).

### Disability

The proportion of participants who described themselves as having a disability was slightly higher in Study 2 (81.4%) than in Study 1 (72.5%).

### Employment Status of Participants

A comparison of the employment status of the participants in Studies 1 and 2 is shown in Figure 21.

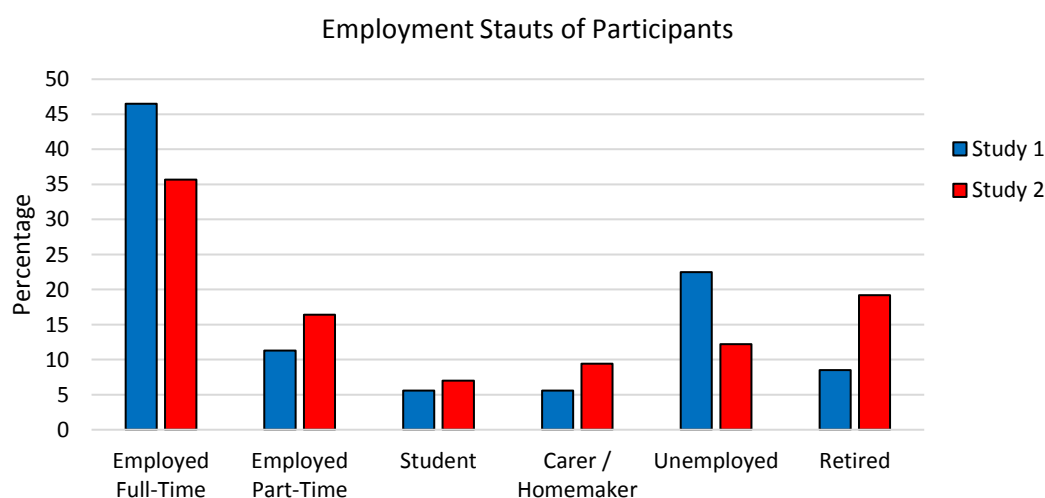


Figure 21: Comparison of the employment status of participants who attended the AYL course in Study 1 and Study 2

### How Participants Heard about the AYL Course

In both studies, the majority of participants heard about the AYL course from their GP (52.8% in Study 2 and 45.2% in Study 1). Figure 22 shows the breakdown of all the sources that provided participants with information about the course for Study 2.

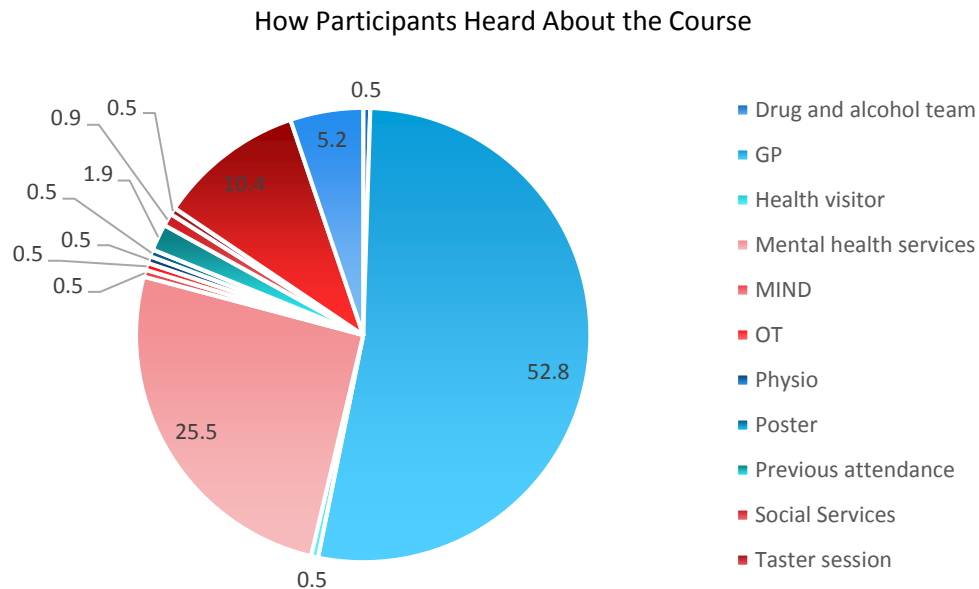


Figure 22: How participants heard about AYL

### Participants' Goals for the Course

Participants were asked to identify their goals for the course in Study 2. These are shown in Figure 23.

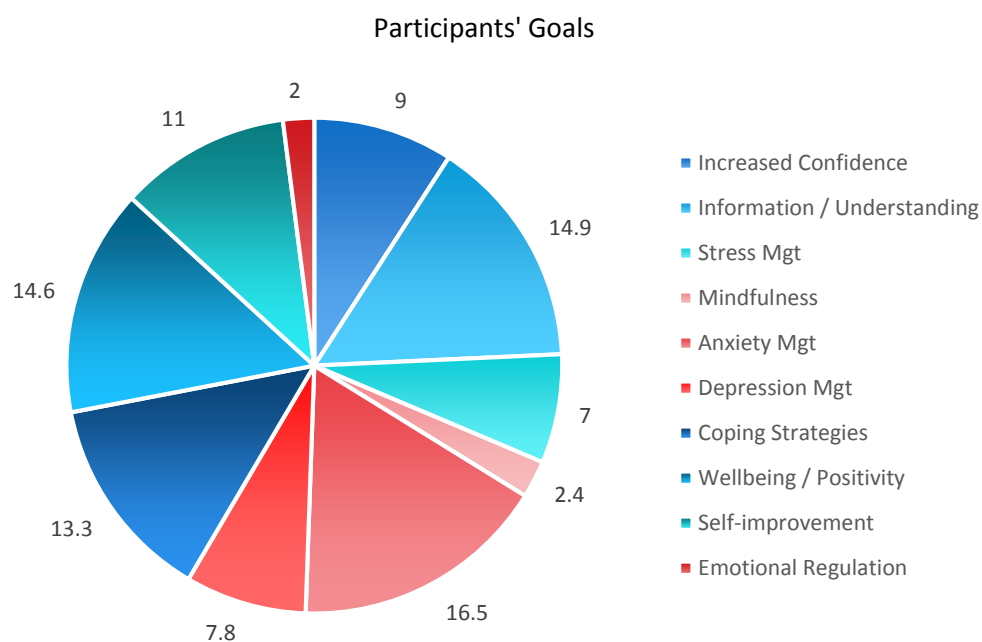


Figure 23: Participants' goals for the course

### Attendance

The attrition rate for the course was 48.9%.

### Pre-Intervention Symptoms

The breakdown of pre-intervention scores by severity range on all outcome measures is shown in Tables 9 - 11. The MBSES-R does not have severity ranges and, therefore, is not reported here.

	Non-Clinical	Mild	Moderate	Moderate-Severe	Severe
GAD-7	12.8% (25)	24.0% (47)	30.6% (60)	N/A	32.7% (64)
PHQ-8	11.2% (22)	17.8% (35)	24.4% (48)	24.4% (48)	22.3% (44)

Table 9: Breakdown of pre-intervention depression and anxiety scores by severity range

	Below Normal Range	Normal Range	Above Normal Range
Rosenberg Self-Esteem Scale	37.8% (74)	62.2% (122)	0.0% (0)

Table 10: Breakdown of pre-intervention self-esteem scores by severity range

	Non-Clinical	Clinical
AAQ-ii	37.5% (69)	62.5% (115)

Table 11: Breakdown of pre-intervention psychological flexibility scores by severity range

### Comparison of Pre-Intervention Scores between Study 1 and Study 2

The pre-intervention scores for depression, anxiety and psychological flexibility in Study 1 and Study 2 were compared. The results are shown in Figures 24 - 26.

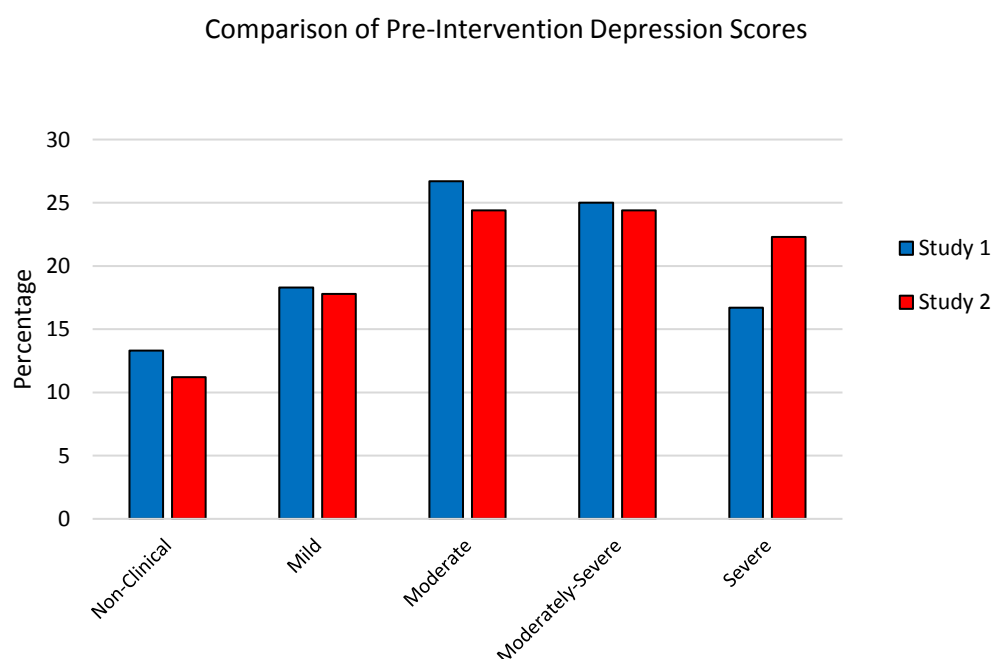


Figure 24: Comparison of pre-intervention depression scores in Study 1 and Study 2

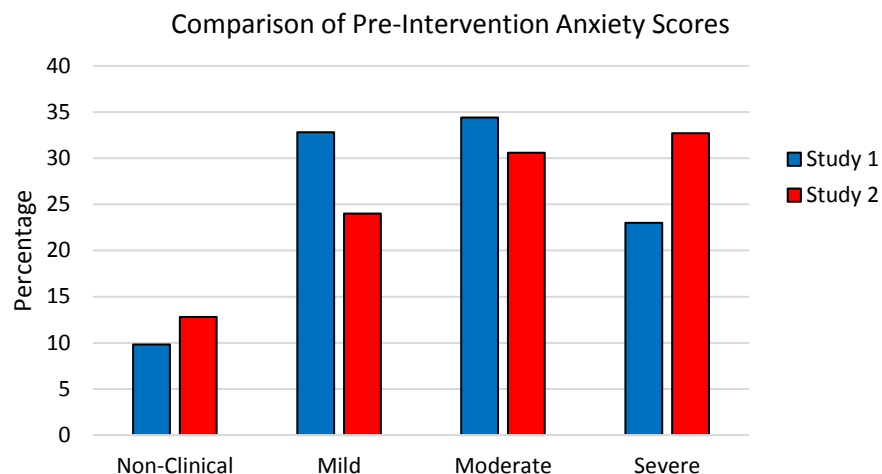


Figure 25: Comparison of pre-intervention anxiety scores for Study 1 and Study 2

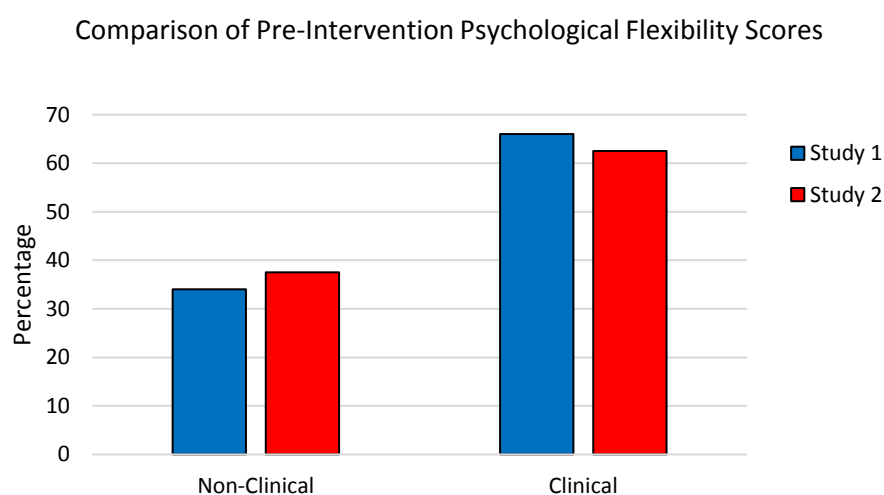


Figure 26: Comparison of pre-intervention psychological flexibility scores for Study 1 and Study 2

Independent t-tests were used to determine whether the pre-intervention anxiety, depression and psychological flexibility scores for Study 1 and Study 2 were significantly different from each other. The results indicated that the pre-intervention scores in the two studies were not statistically significantly different (anxiety:  $t(241) = -.997$ ,  $p = .329$ , ns.; depression:  $t(248) = -.826$ ,  $p = .409$ , ns.; psychological flexibility  $t(248) = .842$ ,  $p = .400$ , ns.).

### Overall Group Outcomes

Table 13 shows the group outcomes for all participants on all measures for the data obtained in Study 2. A comparison of pre- and post-intervention means can be seen in Figure 27.

Paired samples t-tests were conducted to determine the impact of the interventions on each of the measures used in Study 2. The results indicated that there was a statistically significant improvement in depression, anxiety, psychological flexibility and emotional regulation. There was no statistically significant change in self-esteem, distress tolerance or equanimity or overall mindfulness-based self-efficacy.

	N	Pre-Intervention Mean (SD)	Post-Intervention Mean (SD)	Pre-Post Change Mean (SD)	95% CI range	t	Cohen's d	Effect Size
PHQ-8	108	12.73 (5.89)	10.07 (5.45)	-2.43 (4.25)	1.838 – 3.477	6.426**	0.62	Medium
GAD-7	108	10.81 (5.41)	8.55 (5.07)	-2.40 (4.28)	1.405 – 3.114	5.242**	0.50	Medium
RSES	107	15.17 (2.57)	15.43 (2.59)	0.27 (2.86)	0.779 – 0.276	-0.966 (n.s.)	0.09	Small
AAQ-II	106	29.18 (9.98)	26.72 (8.90)	-3.31 (10.03)	0.776 – 4.148	2.896*	0.28	Small
MBSES-R (Emotional Regulation)	107	10.94 (3.82)	12.18 (3.73)	1.10 (3.45)	-1.863 – 0.604	-3.886**	0.38	Small
MBSES-R (Equanimity)	107	7.33 (2.12)	7.20 (1.91)	-0.18 (2.39)	-0.322 – 0.584	0.573 (n.s.)	0.06	Small
MBSES-R (Distress Tolerance)	107	5.39 (1.91)	5.44 (1.61)	-0.05 (2.22)	-0.452 – 0.338	-0.284 (n.s.)	0.03	Small
MBSES-R Total	107	23.64 (5.31)	24.36 (5.73)	0.56 (5.32)	-1.712 – 4.148	-1.477 (n.s.)	0.15	Small

\*  $p < 0.05$ , \*\*  $p < 0.001$

Table 13: Group outcomes on all outcome measures

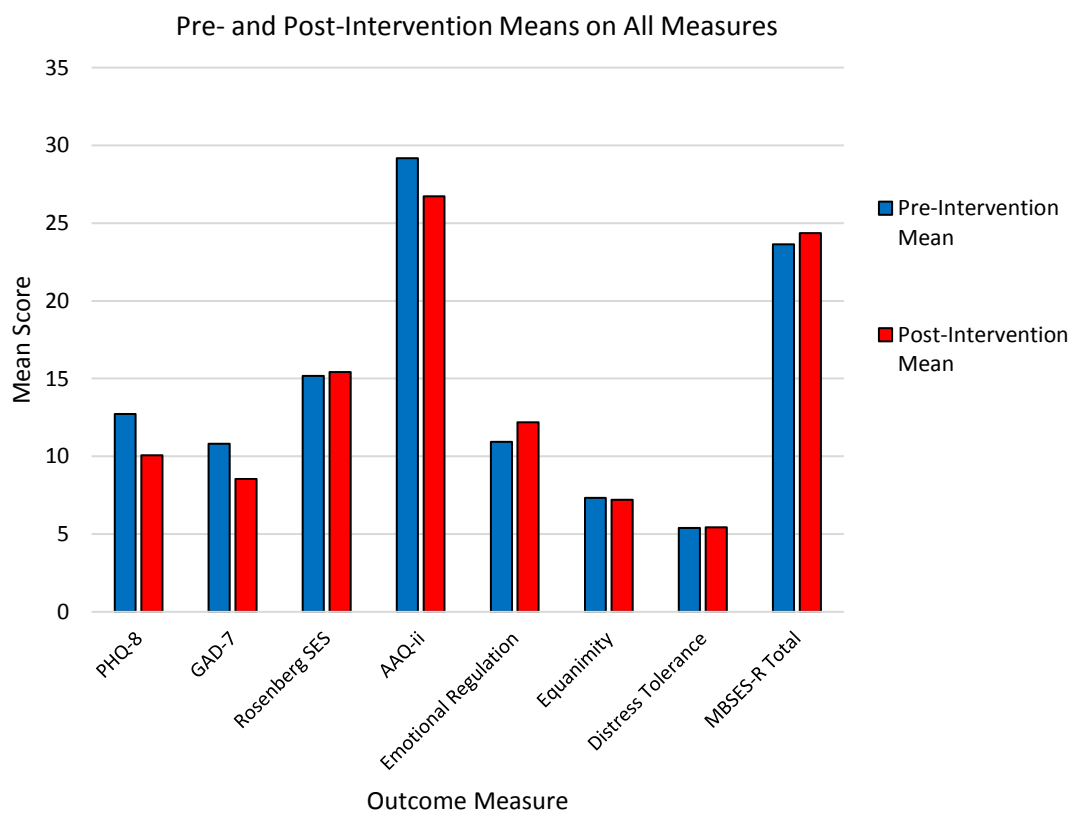


Figure 27: A comparison of pre- and post-intervention means on all measures

### Comparison of Group Outcomes for Study 1 and Study 2

Independent t-tests were used to determine whether the outcomes for depression, anxiety and psychological flexibility were significantly different between Study 1 and Study 2. The results from these showed that the outcomes for depression were significantly different between the two studies. Outcomes for anxiety and psychological flexibility were not significantly different (anxiety:  $t(102.825) = -2.770$ ,  $p < .01$ ; psychological flexibility:  $t(138) = .347$ ,  $p > .05$ , ns). Figure 28 shows a comparison of the group outcomes in Study 1 and Study 2.

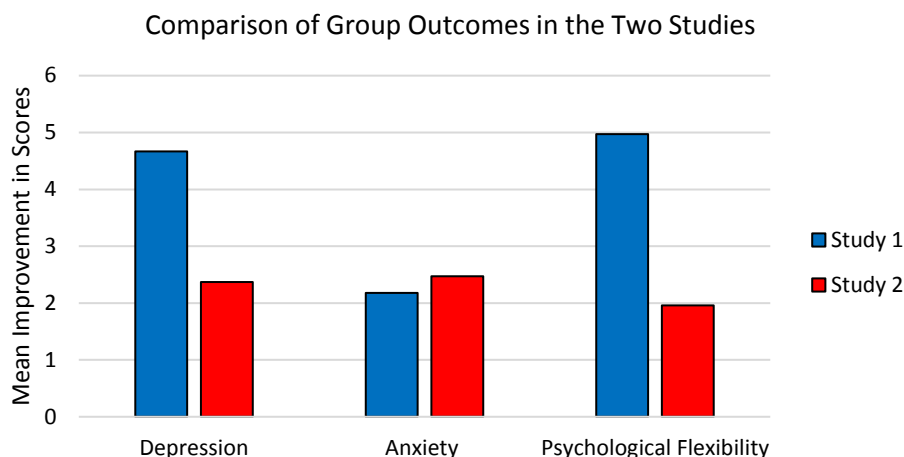


Figure 28: Comparison of group outcomes in Study 1 and Study 2

### Benchmarking

For anxiety and depression, the group outcomes were compared against four benchmarks: the results of Study 1; guided self-help interventions; Cartwright and Hooper (2017); and, the ABUHB evaluation (2019).

#### *Anxiety (GAD-7)*

The benchmark comparisons indicated that Study 2 showed a smaller effect size as measured by the GAD-7 than all of the benchmarks used. These results are shown in Figure 29.

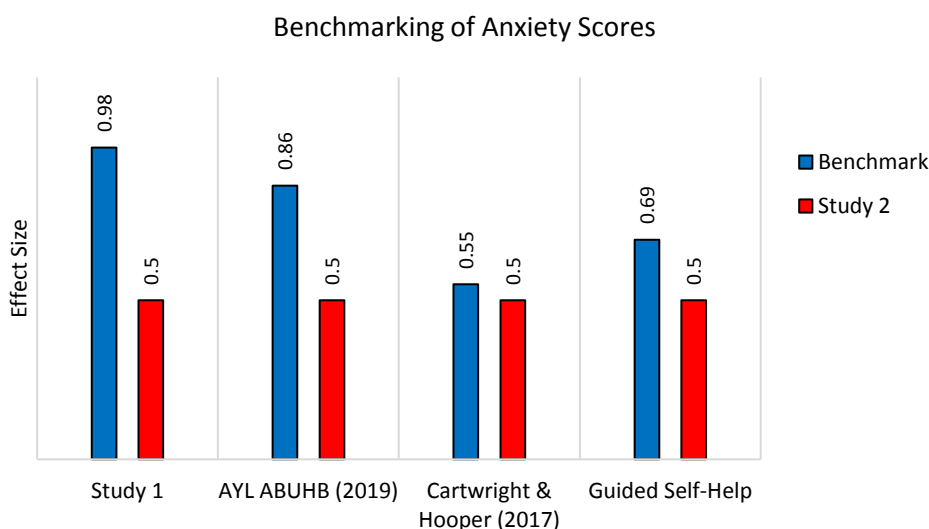


Figure 29: GAD-7 change effect sizes for Study 2 compared to four benchmarks



### Depression (PHQ-9/PHQ-8)

The benchmark comparisons indicated that Study 2 showed a slightly larger effect size for depression than the results reported by Cartwright and Hooper (2017) but smaller effect sizes than the other benchmarks. These results are shown in Figure 30.

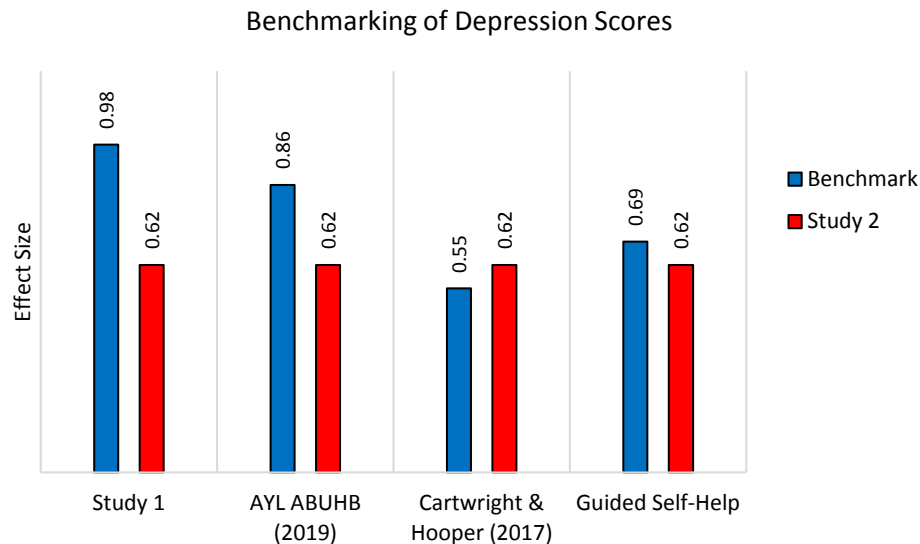


Figure 30: Effect size for depression in Study 2 compared to benchmarks

### Psychological Flexibility (AAQ-ii)

The outcomes for psychological flexibility in Study 2 were compared against three benchmarks: Study 1, Cartwright and Hooper (2017) and a meta-analysis of guided self-help interventions for mindfulness/acceptance (Cavanagh *et al.*, 2014). These comparisons indicated that Study 2 had a similar effect size to Cartwright and Hooper (2017) but a smaller effect size than the other benchmark studies. These results are shown in Figure 31.

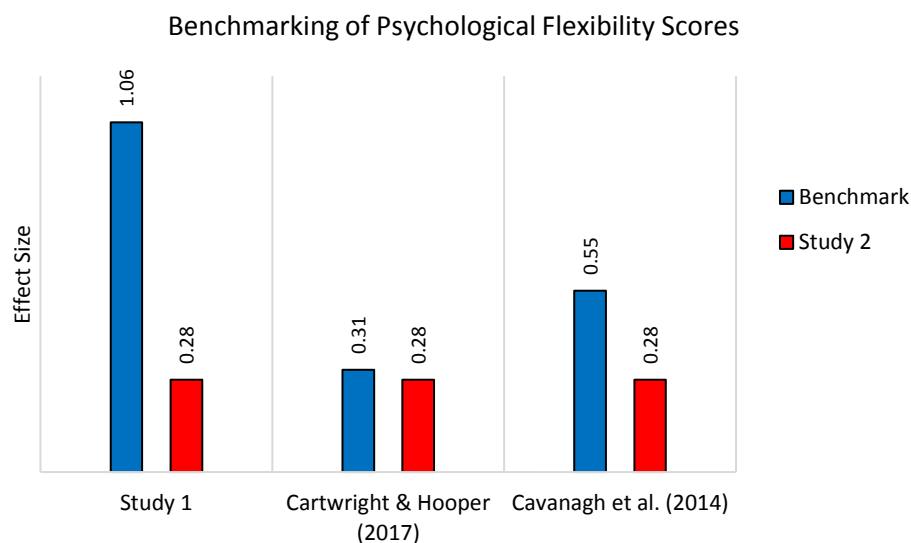


Figure 31: AAQ-ii change effect sizes for Study 2 compared to benchmarks

### Overall Individual Outcomes

Individual outcomes were calculated using a number of measures: clinically reliable change, clinically significant change and MCID (see Table 14). It was not possible to provide this information for the MBSES-R measures as there is no reliability information available for the subscales, therefore, the RCI cannot be calculated.

	N	Positive Clinically Reliable Change	Positive MCID	Reliable Recovery	Stasis	Negative Clinically Reliable Change	Reliable Deterioration	Negative MCID
PHQ-8	102	9 (8.8%)	30 (29.4%)	7 (6.9%)	93 (91.1%)	0 (0.0%)	0 (0.0%)	3 (2.9%)
GAD-7	102	21 (20.6%)	38 (37.3%)	17 (16.7%)	78 (76.5%)	3 (2.9%)	2 (2.0%)	8 (7.8%)
AAQ-ii	102	24 (23.5%)	N/A	13 (12.7%)	71 (69.6%)	7 (6.9%)	4 (3.9%)	N/A
RSES	102	20 (19.6%)	N/A	13 (12.7%)	72 (70.6%)	10 (9.8%)	8 (7.8%)	N/A

Table 14: Individual outcomes for all participants

### Benchmarking

The benchmark provided by the English IAPT services is that 50% of clients should show reliable recovery. For Study 2, this figure is below 50% for anxiety, depression and psychological flexibility scores (see Table 14).

### Comparison of Individual Outcomes in Study 1 and Study 2

Figure 32 shows a comparison of the individual outcomes for depression in Study 1 and Study 2.

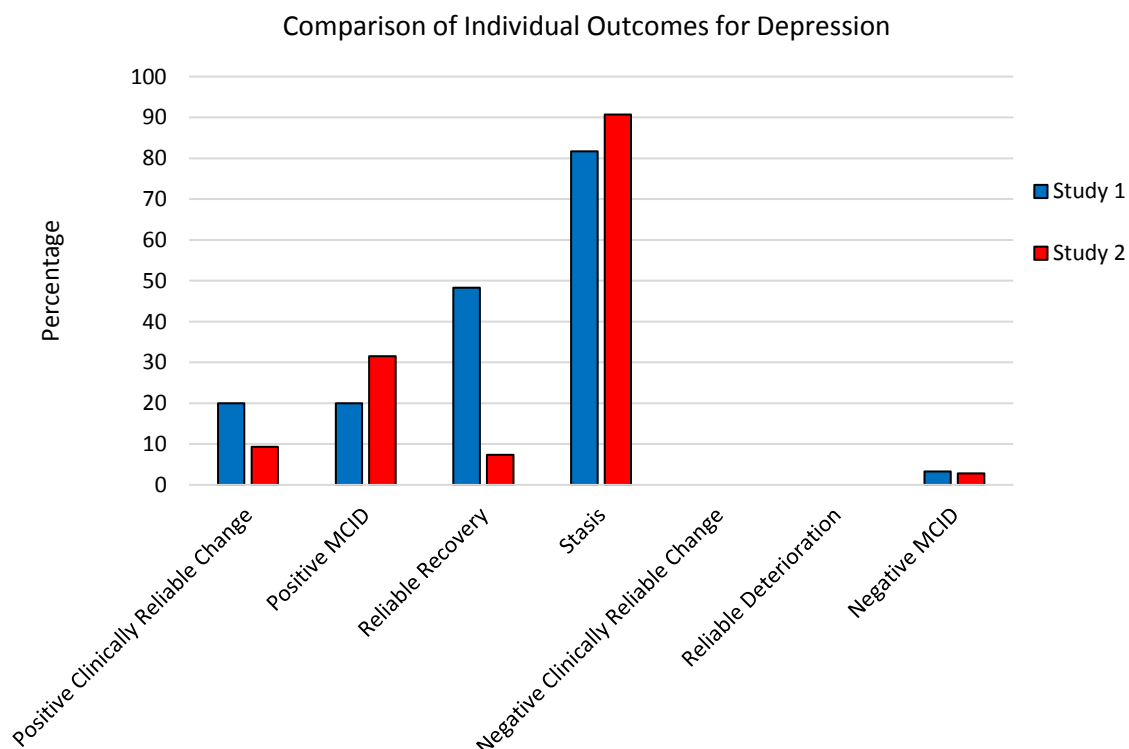
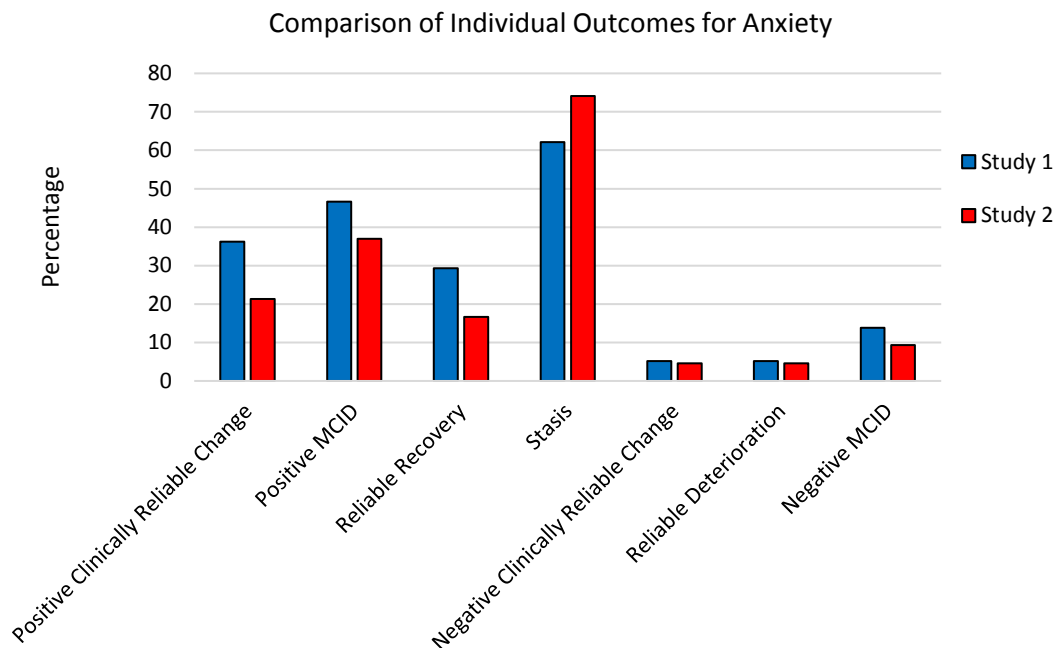


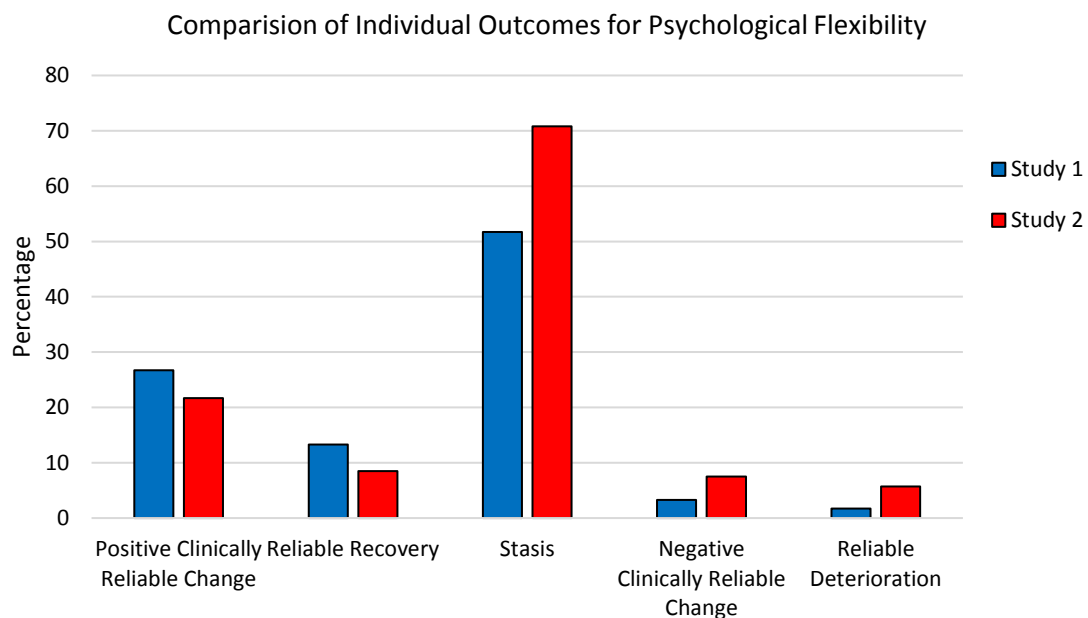
Figure 32: Comparison of the individual outcomes for depression in Study 1 and Study 2

Figure 33 shows a comparison of the individual outcomes for anxiety in Study 1 and Study 2.



*Figure 33: Comparison of the individual outcomes for anxiety in Study 1 and Study 2*

Figure 34 shows a comparison of the individual outcomes for psychological flexibility in Study 1 and Study 2.



*Figure 34: Comparison of the individual outcomes for psychological flexibility Study 1 and Study 2*

### The Impact of Initial Symptom Severity on Outcome

#### *Group Outcomes by Initial Symptom Severity*

Pre- and post-intervention scores were analysed to determine whether initial symptom severity had an impact on the effectiveness of the AYL intervention for anxiety, depression, psychological flexibility and self-esteem. This information is displayed in Table 15.

Severity	N	Pre- Interven- tion Mean (SD)	Post- Interven- tion Mean (SD)	Pre-Post Change Mean (SD)	95% CI range	t	Cohen's d	Effect Size
<b>GAD-7</b>								
Mild - Moderate	66	7.38 (3.66)	6.50 (4.28)	-0.88 (3.99)	-0.10 – 1.86	1.791 (n.s.)	0.22	Small
Moderately Severe – Severe	42	16.19 (2.55)	11.76 (4.55)	-4.43 (4.39)	3.06 – 5.80	6.537* *	1.00	Large
<b>PHQ-9</b>								
Mild - Moderate	62	8.66 (4.03)	7.50 (4.32)	-1.16 (3.82)	0.19 – 2.13	2.393* *	0.30	Small
Moderately Severe – Severe	46	18.22 (2.62)	13.54 (4.89)	-4.67 (4.11)	3.45 – 5.89	7.713* *	1.14	Large
<b>AAQ-II</b>								
Non-Clinical	43	19.35 (6.23)	20.30 (7.91)	0.95 (9.92)	-4.01 – 2.10	-0.630 (n.s.)	- 0.09	Small
Clinical	63	35.89 (5.45)	31.10 (6.60)	-4.79 (7.04)	3.02 – 6.57	5.407* *	0.68	Medium
<b>RSES</b>								
Low	36	12.33 (1.49)	14.14 (2.54)	1.83 (2.99)	-2.802 - -0.809	-3.678* *	0.61	Medium
Normal	71	16.61 (1.63)	16.09 (2.38)	-0.70 (2.34)	-0.044 – 1.087	1.838 (n.s.)	0.22	Small

\* p<0.05 \*\* p<0.001

*Table 15: Group outcomes by initial symptom severity*

The results of independent t-tests on the GAD-7 indicated that individuals who scored as moderately-severe to severely anxious prior to intervention showed a significantly greater reduction in anxiety than those categorised with mild to moderate anxiety ( $t(106) = 4.336$ ,  $p<.001$ ).

The results of independent t-tests on the PHQ-8 showed that individuals who scored as moderately-severe to severely depressed prior to the intervention showed a significantly greater reduction in depression than those categorised with mild to moderate anxiety ( $t(106) = 4.574$ ,  $p<.001$ ).

The results of independent t-tests on the AAQ-II showed that individuals whose degree of psychological inflexibility was in the clinical range prior to the intervention showed a significantly greater reduction in symptoms than those with non-clinical symptoms ( $t(104) = 3.491$ ,  $p<.05$ ).

The results of independent t-tests on the RSES showed that individuals whose self-esteem was in the low range prior to the intervention showed a significantly greater increase in self-esteem than those in the normal range ( $t(105) = 4.818$ ,  $p<.01$ ).

### Individual Outcomes by Initial Symptom Severity

A variety of individual outcome measures were used to determine whether initial symptom severity had an impact in the effectiveness of the AYL intervention for anxiety, depression, psychological flexibility and self-esteem. This information is displayed in Table 16.

	N	Positive Clinically Significant Change	Positive MCID	Reliable Recovery	Stasis	Reliable Deterioration	Negative Clinically Significant Change	Negative MCID
<b>GAD-7</b>								
<b>Mild - Moderate</b>	66	7 (10.6%)	15 (22.7%)	7 (10.6%)	54 (81.8%)	4 (6.1%)	5 (7.6%)	9 (13.6%)
<b>Moderately Severe - Severe</b>	42	16 (38.1%)	25 (60.0%)	11 (26.2%)	26 (61.9%)	0 (0.0%)	0 (0/0%)	1 (2.4%)
<b>PHQ-8</b>								
<b>Mild - Moderate</b>	62	2 (3.2%)	12 (19.4%)	2 (3.2%)	60 (96.8%)	0 (0.0%)	0 (0.0%)	3 (4.8%)
<b>Moderately Severe - Severe</b>	46	8 (17.4%)	22 (47.8%)	6 (13.0%)	38 (82.6%)	0 (0/0%)	0 (0/0%)	0 (0.0%)
<b>AAQ-ii</b>								
<b>Non-Clinical</b>	43	6 (14.0%)	N/A	3 (7.0%)	29 (67.4%)	N/A	8 (18.6%)	N/A
<b>Clinical</b>	63	17 (27.0%)	N/A	6 (9.5%)	46 (73.0%)	0 (0.0%)	0 (0.0%)	N/A
<b>RSES</b>								
<b>Low</b>	36	13 (36.1%)	N/A	15 (41.7%)	19 (52.8%)	N/A	2 (5.6%)	N/A
<b>Normal</b>	71	3 (4.2%)	N/A	3 (4.2%)	58 (81.7%)	3 (4.2%)	9 (12.7%)	N/A

*Table 16: Post-intervention individual outcomes by initial symptom severity*

Chi-square tests were used to determine whether reliable recovery rates were impacted by the participants' initial symptom severity.

Reliable recovery rates for individuals in the moderately-severe to severe anxiety category were significantly different compared to those in the mild to moderate anxiety category ( $\chi^2(1) = 4.488$ ,  $p < .05$ ). See Figure 35.

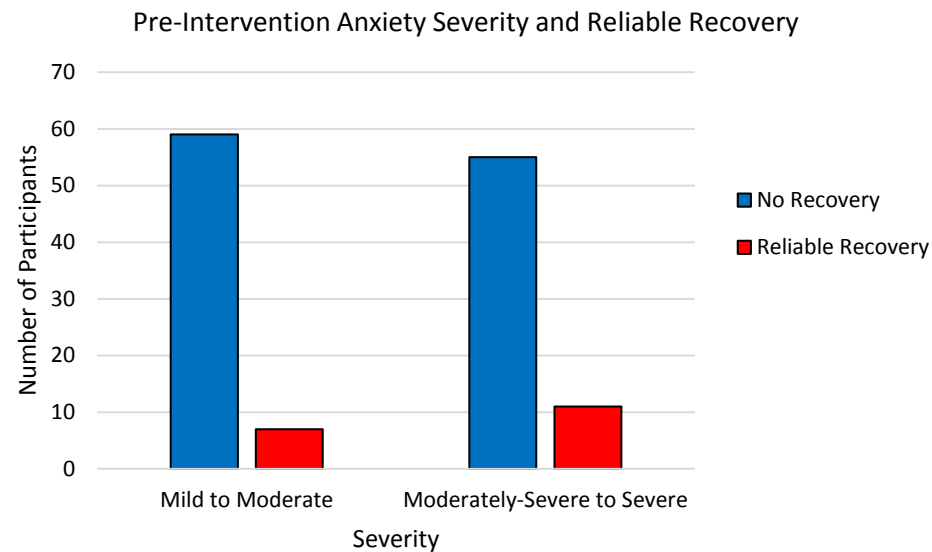


Figure 35: Number of participants who demonstrated reliable recovery calculated on the GAD-7 by initial symptom severity

Differences between reliable recovery rates for individuals with moderately-severe to severe depression were not significantly different to those for individuals in the mild to moderate depression category ( $\chi^2(1) = 3.711$ ,  $p = .054$ , ns). See Figure 36.

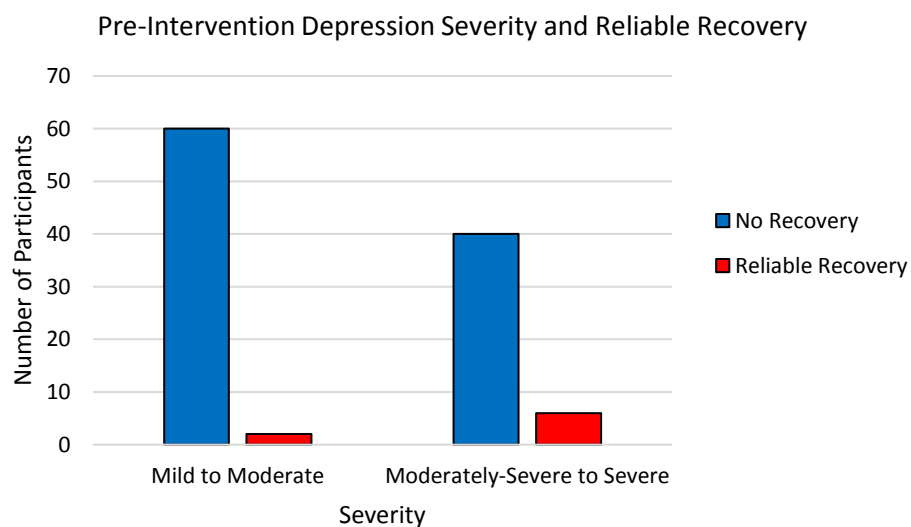
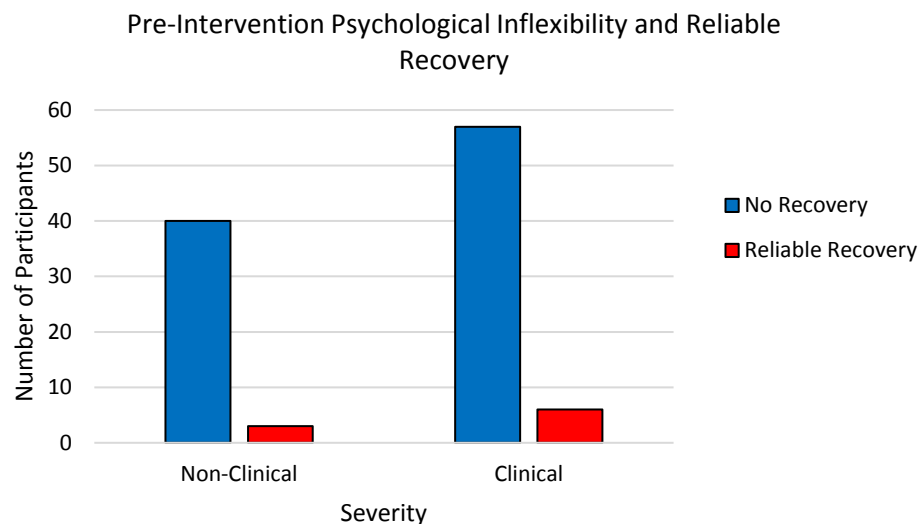


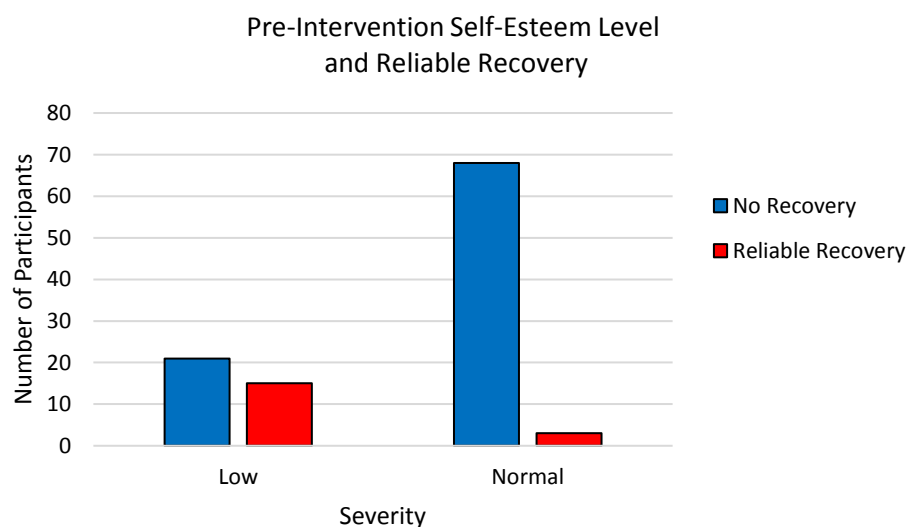
Figure 36: Number of participants who demonstrated reliable recovery on the PHQ-8 calculated by initial symptom severity

Reliable recovery rates for individuals with pre-intervention levels of psychological inflexibility in the clinical range were not significantly different to those for individuals in non-clinical range ( $\chi^2(1) = 0.213$ ,  $p = .664$ ). See Figure 37.



*Figure 37: Number of participants who demonstrated reliable recovery on the AAQ-II calculated by initial symptom severity*

Reliable recovery rates for individuals with pre-intervention levels of self-esteem in the low range were significantly different to those for individuals in normal range ( $\chi^2(1) = 23.932, p < .01$ ). See Figure 38.



*Figure 38: Number of participants who demonstrated reliable recovery on the RSES calculated by initial symptom severity*

### Feedback from Participants

98.3% of attendees who responded stated that they would recommend the AYL course to others. This is slightly higher than for attendees of Study 1, 95.6% of whom stated that they would recommend the course.

Participants were asked to indicate how useful they found the course overall. These results can be seen in Figure 39.

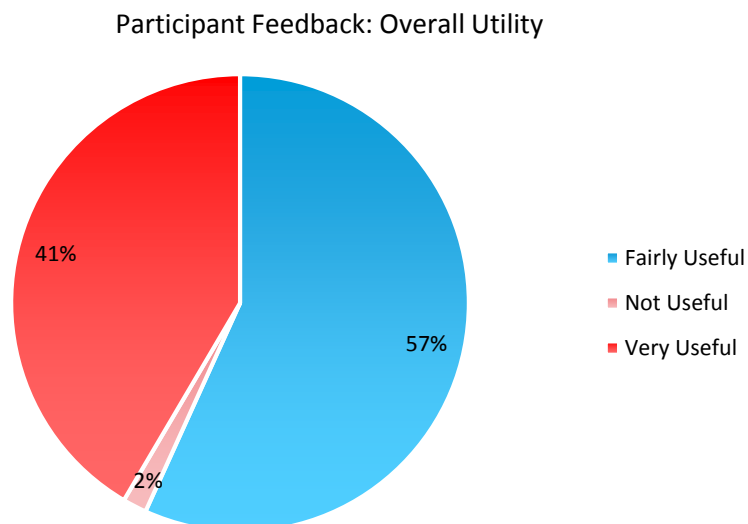


Figure 39: Participants' feedback regarding the utility of the course

Figure 40 shows the comparison between the participant view of the overall utility of the course in Study 1 and Study 2. A chi-square test indicated that the overall utility of the course was rated as being statistically significantly better in Study 2 than in Study 1 ( $\chi^2(3) = 9.570, p < .05$ ).

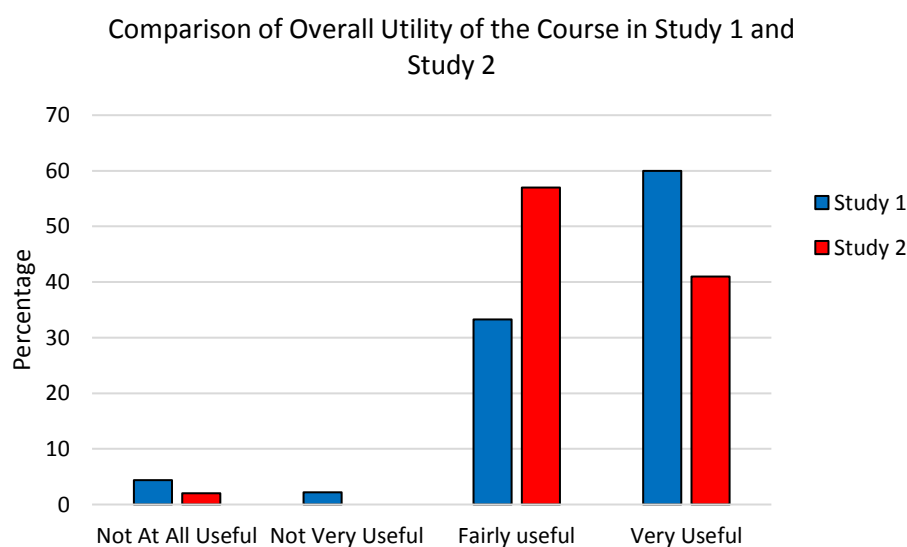


Figure 40: Comparison of the participants' perception of the overall utility of the AYL course in Study 1 and Study 2.

Participants were asked to rate how well they felt the course covered the main factors that AYL is designed to include on a 5-point Likert scale (from "Not at all" to "A lot"). The results are shown in Table 17.

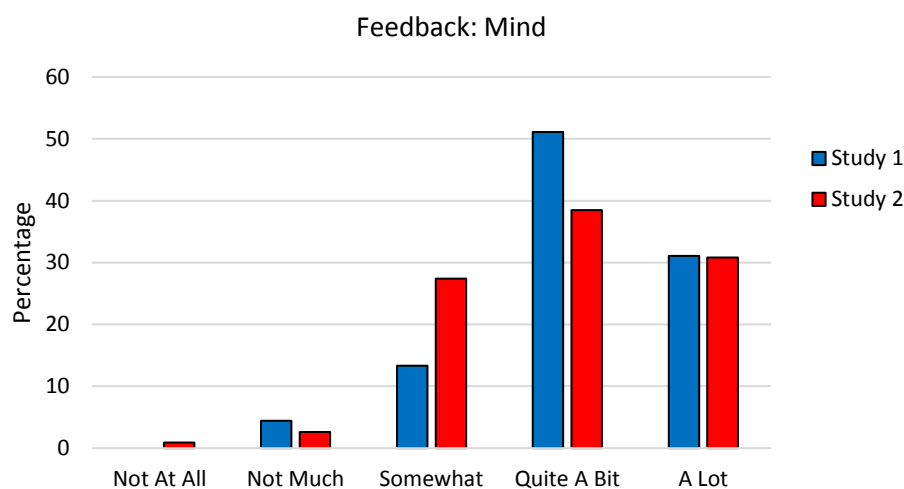


Course Focus	Not at all	Not much	Somewhat	Quite a bit	A lot
Mind	0.9%	2.6%	27.4%	38.5%	30.8%
Values	2.6%	7.8%	34.8%	33.9%	20.9%
Thoughts	0.9%	5.3%	26.5%	35.4%	31.9%
Actions	0.9%	10.5%	28.1%	37.7%	22.8%

*Table 17: Participants' feedback regarding the coverage of the main aspects of the AYL course*

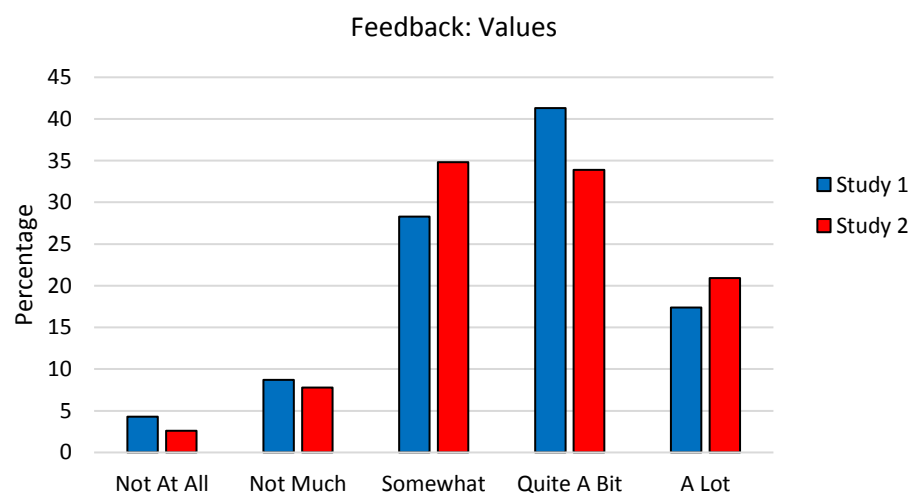
This feedback was compared to the feedback obtained from participants in Study 1. Chi-square tests were used to determine whether the feedback obtained from participants in Study 1 and Study 2 were statistically significant.

The feedback for the Mind factor was not significantly different between Study 1 and Study 2 ( $\chi^2(4) = 4.719$ ,  $p = .317$ ). See Figure 41.



*Figure 41: Comparison between participant feedback for understanding of mind obtained on AYL course for Study 1 and Study 2*

The feedback for the Values factor was not significantly different between Study 1 and Study 2 ( $\chi^2(4) = 1.820$ ,  $p = .769$ ). See Figure 42.



*Figure 42: Comparison between participant feedback for understanding of values obtained on AYL course for Study 1 and Study 2*

The feedback for the Values factor was not significantly different for Study 1 and Study 2 ( $\chi^2(4) = 2.539$ ,  $p = .638$ ). See Figure 43.

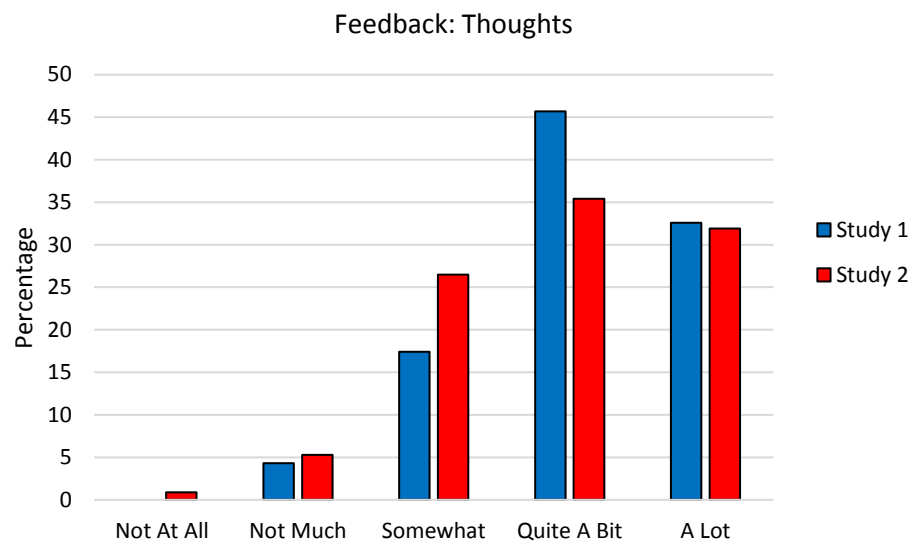


Figure 43: Comparison between participant feedback for understanding of thoughts obtained on AYL course for Study 1 and Study 2

The feedback for the Actions factor was not significantly different between Study 1 and Study 2 ( $\chi^2(4) = 1.533$ ,  $p = .821$ ). See Figure 44.

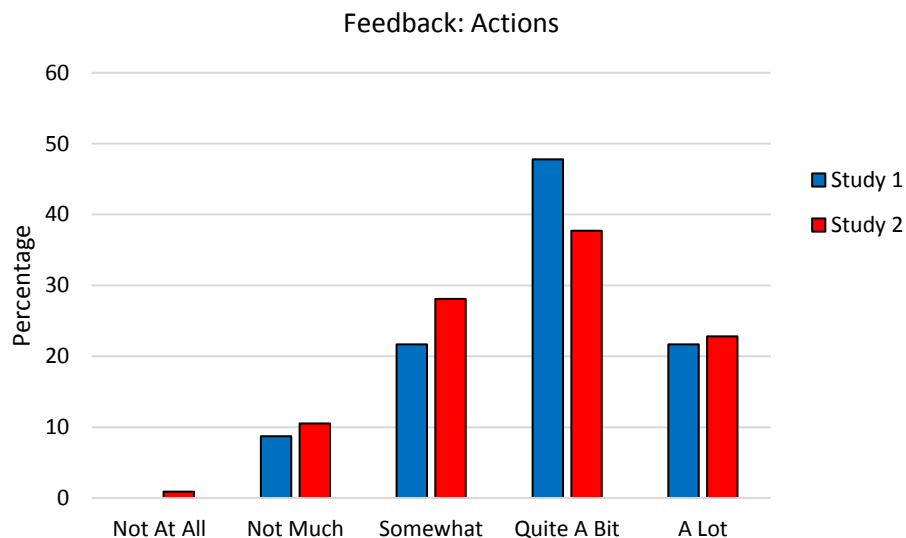


Figure 44: Comparison between participant feedback for understanding of actions obtained on AYL course for Study 1 and Study 2

## Discussion

### Summary of Results

When comparing the participant demographics from Study 1 and Study 2, it can be seen that Study 2 had a less equal proportion of male and female participants than Study 1. The ethnicity of participants

was approximately equal in Study 1 and Study 2, although Study 1 had a wider range of ethnicities. This may be a reflection of the fact that Study 2 included data from only one county, whereas Study 1 included data from three counties which is likely to have increased the ethnic diversity of the participants. In both studies, the majority of the participants (approximately half) were referred to AYL via their GPs.

Pre-intervention anxiety, depression and psychological flexibility scores for Study 1 and Study 2 were not significantly different from each other.

Group outcomes in Study 2 showed that there was a statistically significant improvement in depression, anxiety, psychological flexibility and emotional regulation. There was no statistically significant change in self-esteem, distress tolerance, equanimity or overall mindfulness-based self-efficacy. Comparing the results of Study 2 with the benchmark studies showed that there was a smaller effect size for anxiety outcomes compared to the results of Study 1 and all other benchmarks apart from the Cartwright and Hooper (2017) study where the results were approximately similar. When considering depression outcomes, the effect size found in Study 2 was approximately comparable to the Cartwright and Hooper (2017) study and the results from the meta-analysis of guided self-help interventions (Coull & Morris, 2011) but smaller than that reported in Study 1 and ABUHB (2019). The changes in psychological flexibility in Study 2 were very small. The effect size was approximately the same as that reported by Cartwright and Hooper (2017), but less than that reported by Cavanagh *et al.* (2014) and the results found in Study 1. Overall, comparisons of effect sizes for between Study 1 and Study 2 indicated that Study 1 was more effective than Study 2 on all outcome measures. The effect size for anxiety symptoms in Study 1 was nearly twice as big as in Study 2. For depression, the effect size in Study 1 was approximately a third larger than in Study 2. For psychological flexibility, the effect size in Study 1 was nearly five times larger than in Study 2.

When looking at data on an individual level, the majority of participants did not show any reliable recovery on any of the outcome measures used. For depression, reliable recovery rates were 6.9%; for anxiety, they were 16.7%; and, for both psychological flexibility and self-esteem, reliable recovery rates were 12.7%. This is well below the benchmark figure of 50% from IAPT England. The majority of the participants fell in the stasis (no change) category on all outcome measures. It should be noted that only a small number of participants showed any clinical deterioration, suggesting that the intervention did not cause any harm. Comparisons of individual outcomes between Study 1 and Study 2 showed that, despite both studies being below the IAPT England benchmark, the results from Study 1 were more positive on all outcome measures. This was particularly the case with depressive symptoms where 48.3% of the participants in Study 1 showed reliable recovery.

Both the group and individual outcome data indicated that participants with more severe initial symptoms showed significantly greater improvement than those with milder initial symptoms. For group outcome measures, individuals with moderately-severe to severe initial symptoms showed a statistically significant ( $p < .01$ ) improvement with a large effect size for all outcome measures. Individuals with mild to moderate initial symptoms, showed no significant improvement for anxiety or psychological flexibility and only a small effect size on all outcome measures. In terms of the individual data, this difference only reached statistical significance for the GAD-7.

The attrition rate of participants in Study 2 was significantly lower than for Study 1 (48.9% compared to 13.2%). It should be noted that in Study 2 this figure assumes that participants attended all four sessions if they completed both pre- and post-intervention measures. This assumption is likely to lead to an under-estimation of the attrition rate in Study 2.

The qualitative feedback from participants indicated that over 98% of attendees would recommend the course to others. This was slightly higher than for Study 1. The overall utility of the intervention in Study 2 was rated by participants as being significantly higher than in Study 1. The participant feedback regarding how well the course covered each of the main topics of AYL was not significantly different between Study 1 and Study 2.

#### Limitations of the Evaluation

Study 2 suffers from the same limitations as described in Study 1. In addition, Study 2 only looked at pre- and post-intervention data, rather than session by session data. This has resulted in less detailed information about the effectiveness of course over time and the attrition rates of the participants.

Study 2 used three subscales of MBSES-R to look at changes in some aspects of mindfulness-based self-efficacy. It is unclear why this outcome measure was chosen and why only half of the subscales were used. Whilst it is useful to have a measure of mindfulness-based self-efficacy given the emphasis on mindfulness in the AYL course, it would have been more appropriate to use a measure that has had its psychometric properties more thoroughly investigated. The lack of reliability data for the MBSES-R makes its inclusion in this study questionable and prevented some of statistical analyses being carried out.

#### Conclusions

The differences between the clinical outcomes and participant feedback in the two studies suggest that there may be differences in the way in which the course is delivered in the three counties and/or how it has been delivered over the past few years. It should also be noted that the interventions in Study 2 was more highly rated by participants than the interventions in Study 1 despite the latter having better clinical outcomes.

#### Recommendations

- HDUHB should use the most up-to-date version of the AYL course as it is possible that the version of the course used was one of the reasons for the different clinical outcomes found in Study 1 and Study 2.
- LPMHSS staff should be provided with training on facilitating AYL as this would increase their confidence and would also be in-line with the Matrics Cymru recommendation that staff should be properly trained in the interventions that they provide. Staff training will also ensure consistency of delivery across sites in the HDUHB area.
- Strategies to maintain engagement with clients at risk of dropping out of AYL need to be developed and evaluated.
- AYL is not currently being offered in Welsh, which goes against the legally-binding Welsh Language Standards (Welsh Government, 2018). Whilst it is recognised that the current staff mix in LPMHSS means that it is difficult for AYL to be delivered in Welsh, creative solutions should be developed to address this difficulty. For example, the use of interpreters; translating course materials into Welsh or providing a video-recorded version of the course in Welsh that could be provided to Welsh-speaking participants to view in their own time (subject to copyright).
- Consideration should be given to allowing more expensive community venues to be used for the delivery of AYL and other psychoeducational courses. Given the positive impact that high quality delivery of such courses could have on the LPMHSS waiting list and referrals to other mental health services in HDUHB, the increase in cost is likely to be minimal or non-existent in the long-term.

- AYL should be routinely evaluated on a session-by-session basis to ensure that the effectiveness of the intervention is maintained over time and that outcomes are in-line with those expected. Thought should be given to the questionnaires that are used to ensure that they provide usable data, whilst also not placing too great a burden on the participants who are completing them.
- AYL participants should be offered the option of a telephone review following the intervention to determine whether they need any extra support. It is suggested that this is offered one to two months after the end of the course in order to allow participants the time to consolidate what they have learnt during the intervention before determining whether they require further interventions. It is recommended that this should be offered to all participants even if they dropped-out of the intervention. The information gathered at these reviews could be used to determine the medium-term effects of AYL; gain a better understanding of reasons why participants do not attend the full AYL course; and, monitor the proportion of participants that require further intervention. This review should be standardised and information recorded on a central database to enable comparisons between counties and to facilitate routine audits. This will involve developing ways to accurately record who attends the sessions. One aspect to consider is whether it is necessary, or appropriate, to create a mental health service record on Care Partner for all individuals who attend the course or whether this information should be stored in a different manner.
- During the pandemic, AYL has been offered by Public Health Wales as a series of four pre-recorded videos that can be accessed online. Whilst this type of innovation is a good idea, it should not take fully take the place of face-to-face courses as the normalisation process of being with other people with similar difficulties is something that participants have identified as being helpful. Offering face-to-face courses also ensures that people who are not confident in using the internet or those who are unable to due to financial constraints or lack of connectivity still are able to benefit from psychoeducational courses.
- AYL has a number of different versions including for students, cancer survivors and stroke survivors. Neil Frude has confirmed that the training for the generic and specific versions of AYL is the same. Given this, it may be worth negotiating with other services and organisations so that the cost of training can be spread. For example, oncology and neurology services, clinical health psychology and local universities. Historically AYL was piloted for first years at Aberystwyth University. It is unclear what the results of this pilot were but, given the current focus on student mental health, it may be worth discussing a collaboration with the university. This is likely to reduce pressure on LPMHSS services, particularly in Ceredigion as they routinely see an increase in referrals from university students at the start of the academic year.
- Given LPMHSS staff feedback, it would be important to clarify with GPs that the majority of referrals to LPMHSS are likely to be offered psychoeducational courses in the first instance. This will create more of a focus on high volume, low intensity interventions in LPMHSS which will have a positive impact on waiting lists. Moreover, this will ensure that clients are clear about the care pathway.

### *Acknowledgements*

The author would like to express gratitude to all practitioners within LPMHSS services in HDUHB and to the AYL participants whose engagement and participation made this evaluation possible.

## References

- Addis, M.E. & Jacobson, N.S. (1996). Reasons for depression and the process and outcome of cognitive-behavioral psychotherapies. *Journal of Consulting and Clinical Psychology*, 64(6), 1417 – 1424.
- Aneurin Bevan University Health Board (October 2019) *ACTivate Your Life (AYL) Evaluation*. Personal communication with Neil Frude.
- Arch, J.J., Eifert, G.H., Davies, C., Vilardaga, J.C.P., Rose, R.D. & Craske, M.G. (2012). Randomized clinical trial of cognitive behavioral therapy (CBT) versus acceptance and commitment therapy (ACT) for mixed anxiety disorders. *Journal of consulting and clinical psychology*, 80(5), 750 - 765.
- Bach, P. & Hayes, S. C. (2002). The use of acceptance and commitment therapy to prevent the rehospitalization of psychotic patients: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 70(5), 1129 - 1139.
- Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., ... & Evans, C. (2013). The CORE-10: A short measure of psychological distress for routine use in the psychological therapies. *Counselling and Psychotherapy Research*, 13(1), 3 - 13.
- Baumeister, R.F., Campbell, J.D., Krueger, J.I. & Vohs, K.D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, 4(1), 1 - 44.
- Beard, C. & Björgvinsson, T. (2014). Beyond generalized anxiety disorder: Psychometric properties of the GAD-7 in a heterogeneous psychiatric sample. *Journal of Anxiety Disorders*, 28, 547 - 552.
- Beard, C., Hsu, K., Rifkin, L., Busch, A. & Björgvinsson, T. (2016). Validation of the PHQ-9 in a psychiatric sample. *Journal of Affective Disorders*, 193, 267 - 273.
- Bond, F.W., Hayes, S.C., Baer, R.A., Carpenter, K.M., Guenole, N., Orcutt, H.K., Waltz, T. & Zettle, R. D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire – II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy*, 42, 676 – 688.
- Borek, A.J. & Abraham, C. (2018). How do small groups promote behaviour change? An integrative conceptual review of explanatory mechanisms. *Applied Psychological Health Well-Being*, 10, 30 - 61.
- Burlingame, G.M., Strauss, B. & Joyce, A.S. (2013). Change mechanisms and effectiveness of small group treatments. In: M.J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (6th ed.). New Jersey: John Wiley & Sons Inc.
- Byrne, B.M. (1996). *Measuring Self-Concept Across the Lifespan: Issues and Instrumentation*. Washington, DC: American Psychological Association.
- Cartwright, J. & Hooper, N. (2017). Evaluating a transdiagnostic Acceptance and Commitment Therapy psychoeducation intervention. *Cognitive Behaviour Therapist*, 10(e9), 1 - 6.
- Cavanagh, K., Strauss, C., Forder, L. & Jones, F. (2014). Can mindfulness and acceptance be learnt by self-help?: A systematic review and meta-analysis of mindfulness and acceptance-based self-help interventions. *Clinical Psychology Review*, 34(2), 118 - 129.
- Cayoun, B.A., Francis, S.E., Kasselis, N. & Skilbeck, C. (2012). *Mindfulness-Based Self Efficacy Scale-Revised (MSES-R)*.
- Connell, J. & Barkham, M. (2007). *CORE-10 User Manual, Version 1.1*. CORE System Trust & CORE Information Management Systems Ltd.
- Cook, C.E. (2008). Clinimetrics Corner: The Minimal Clinically Important Change score (MCID): A necessary pretense. *Journal of Manual and Manipulative Therapy*, 16(4), E82 - E83.

- Coull, G. & Morris, P.G. (2011). The clinical effectiveness of CBT-based guided self-help interventions for anxiety and depressive disorders: A systematic review. *Psychological Medicine*, 41, 2239 - 2252.
- Dahl, J.C., K.G. Wilson & C. Luciano (2005). *Acceptance and Commitment Therapy for Chronic Pain*. Reno, NV: Context Press.
- Davis, R.N. & Nolen-Hoeksema, S. (2000). Cognitive inflexibility among ruminators and non-ruminators. *Cognitive Therapy and Research*, 24(6), 699 - 711.
- Delgadillo, J., Kellett, S., Ali, S., McMillan, D., Barkham, M., Saxon, D., ... Lucock, M. (2016). A multi-service practice research network study of large group psychoeducational cognitive behavioural therapy. *Behaviour Research and Therapy*, 87, 155 - 161.
- Delgadillo, J., McMillan, D., Leach, C., Lucock, M., Gilbody, S. & Wood, N. (2014). Benchmarking routine psychological services: a discussion of challenges and methods. *Behavioural and Cognitive Psychotherapy*, 42(1), 16 - 30.
- Elliot, A.J., Sheldon, K.M. & Church, M.A. (1997). Avoidance personal goals and subjective well-being. *Personality and Social Psychology Bulletin*, 23(9), 915 - 927.
- Evans, C., Margison, F. & Barkham, M. (1998). The contribution of reliable and clinically significant change methods to evidenced-based mental health. *Evidenced Based Mental Health*, 1, 70 - 72.
- Francis, S. & Cayoun, B.A. (2011). *Measuring mindfulness skills in clinical populations: A preliminary study of the Mindfulness-based Self-Efficacy Scale*. Unpublished manuscript.
- Gifford, E.V., Kohlenberg, B.S., Hayes, S.C., Antonuccio, D.O., Piasecki, M.M., Rasmussen-Hall, M.L. & Palm, K.M. (2004). Acceptance-based treatment for smoking cessation. *Behavior Therapy*, 35(4), 689 - 705.
- Gloster, A.T., Walder, N., Levin, M.E., Twohig, M.P. & Karekla, M. (2020). The empirical status of acceptance and commitment therapy: A review of meta-analyses. *Journal of Contextual Behavioral Science*, 18, 181 – 192.
- Greco, L.A., Lambert, W. & Baer, R.A. (2008). Psychological inflexibility in childhood and adolescence: Development and evaluation of the Avoidance and Fusion Questionnaire for Youth. *Psychological Assessment*, 20(2), 93 – 102.
- Greenberg, J., Solomon, S., Pyszczynski, T., Rosenblatt, A., Burling, J., Lyon, D. *et al.* (1992). Why do people need self-esteem? Converging evidence that self-esteem serves an anxiety-buffering function. *Journal of Personality and Social Psychology*, 63, 913 - 922.
- Gyani, A., Shafran, R., Layard, R. & Clark, D.M. (2013). Enhancing recovery rates: Lessons from year one of IAPT. *Behaviour Research and Therapy*, 51, 597 - 606.
- Hayes, S.C. (1984). Making sense of spirituality. *Behaviorism*, 99 - 110.
- Hayes, S.C., Luoma, J.B., Bond, F.W., Masuda, A. & Lillis, J. (2006). Acceptance and Commitment Therapy: Model, Processes and Outcomes. *Behaviour Research and Therapy*, 44(1), 1 - 25.
- Hayes, S.C. & Strosahl, K.D. (Eds.). (2004). *A Practical Guide to Acceptance and Commitment Therapy*. New York: Springer-Verlag.
- Hayes, S.C., Strosahl, K.D. & Wilson, K.G. (1999). *Acceptance and Commitment Therapy: An Experiential Approach to Behavior Change*. Guilford Press.

Hayes, S.C., Wilson, K.G., Gifford, E.V., Follette, V.M. & Strosahl, K. (1996). Experiential avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. *Journal of Consulting and Clinical Psychology*, 64(6), 1152 – 1168.

Holman, E.A. & Silver, R.C. (1998). Getting "stuck" in the past: Temporal orientation and coping with trauma. *Journal of Personality and Social Psychology*, 74(5), 1146 - 1163.

IBM Corp. (2017). *IBM SPSS Statistics for Windows*, Version 25.0. Armonk, NY: IBM Corp.

Jacobson, N.S. & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12 – 19.

Kellett, S., Clarke, S. & Matthews, L. (2007). Delivering group psychoeducational CBT in primary care: Comparing outcomes with individual CBT and individual psychodynamic-interpersonal psychotherapy. *British Journal of Clinical Psychology*, 46(2), 211 - 222.

Kemani, M.K., Olsson, G.L., Lekander, M., Hesser, H., Andersson, E. & Wicksell, R.K. (2015). Efficacy and cost-effectiveness of acceptance and commitment therapy and applied relaxation for longstanding pain. *The Clinical Journal of Pain*, 31(11), 1004 - 1016.

Kertz, S., Bigda-Peyton, J. & Bjorgvinsson, T. (2013). Validity of the Generalized Anxiety Disorder-7 scale in an acute psychiatric sample. *Clinical Psychology and Psychotherapy*, 20(5), 456 - 464.

Kroenke, K., Spitzer, R.L. & Williams, J.B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606 - 613.

Kroenke, K., Strine, T.W., Spitzer, R.L., Williams, J.B., Berry, J.T. & Mokdad, A.H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, 114(1-3), 163 - 173.

Levitt, J.T., Brown, T.A., Orsillo, S.M. & Barlow, D.H. (2004). The effects of acceptance versus suppression of emotion on subjective and psychophysiological response to carbon dioxide challenge in patients with panic disorder. *Behavior Therapy*, 35(4), 747 - 766.

Lightsey, O.R., Burke, M., Ervin, A., Henderson, D. & Yee, C. (2006). Generalized self-efficacy, self-esteem and negative affect. *Canadian Journal of Behavioural Science*, 38, 72 - 80.

Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W. & Herzberg, P.Y. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Medical Care*, 266 - 274.

Löwe, B., Unützer, J., Callahan, C., Perkins, A.J. & Kroenke, K. (2004). Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. *Medical Care*, 42(12), 1194 - 1201.

Lueger, R.J. & Barkham, M. (2010). Using benchmarks and benchmarking to improve quality of practice and services. In: M. Barkham, G.E. Hardy & J. Mellor-Clark (Eds.), *Developing and Delivering Practice-Based Evidence: A Guide for the Psychological Therapies* (p. 223 - 256). Wiley-Blackwell.

Lukens, E.P. & McFarlane, W.R. (2004). Psychoeducation as evidence-based practice: Considerations for practice, research and policy. *Brief Treatment and Crisis Intervention*, 4, 205 - 225.

McCracken, L.M., Sato, A. & Taylor, G. J. (2013). A trial of a brief group-based form of Acceptance and Commitment Therapy (ACT) for chronic pain in general practice: Pilot outcome and process results. *The Journal of Pain*, 14(11), 1398 - 1406.

McMillan, D., Gilbody, S. & Richards, D. (2010). Defining successful treatment outcome in depression using the PHQ-9: A comparison of methods. *Journal of Affective Disorders*, 127, 122 - 129.

Mendolia, M. & Baker, G.A. (2008). Attentional mechanisms associated with repressive distancing. *Journal of Research in Personality*, 42(3), 546 - 563.



Minami, T., Serlin, R.C., Wampold, B.E., Kircher, J.C. & Brown, G.S. (2008). Using clinical trials to benchmark effects produced in clinical practice. *Quality and Quantity*, 42, 513 - 525.

Mohabbat-Bahar, S., Maleki-Rizi, F., Akbari, M.E. & Moradi-Joo, M. (2015). Effectiveness of group training based on Acceptance and Commitment Therapy on anxiety and depression of women with breast cancer. *Iranian Journal of Cancer Prevention*, 8(2), 71 - 76.

National Collaborating Centre for Mental Health (2020). *The Improving Access to Psychological Therapies Manual* (version 4). Available at: <https://www.england.nhs.uk/wp-content/uploads/2020/05/iapt-manual-v4.pdf> [Accessed 9th July 2020].

National Institute for Health and Care Excellence (2011). *Common Mental Health Problems: Identification and Pathways to Care* (NICE Clinical Guideline CG123). Available at: <https://www.nice.org.uk/guidance/cg123/chapter/1-guidance> [Accessed 2<sup>nd</sup> November 2020].

Neustadt, E., Chamorro-Premuzic, T. & Furnham, A. (2006). The relationship between personality traits, self-esteem and attachment at work. *Journal of Individual Differences*, 27, 208 - 217.

Nordin, L. & Rorsman, I. (2012). Cognitive behavioural therapy in multiple sclerosis: A randomized controlled pilot study of acceptance and commitment therapy. *Journal of Rehabilitation Medicine*, 44(1), 87 - 90.

Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. NJ: Princeton University Press.

Schütze, R., Rees, C., Preece, M. & Schütze, M. (2010). Low mindfulness predicts pain catastrophizing in a fear-avoidance model of chronic pain. *Pain*, 148(1), 120 - 127.

Sheldon, K.M. & Elliot, A.J. (1999). Goal striving, need satisfaction and longitudinal well-being: The self-concordance model. *Journal of Personality and Social Psychology*, 76(3), 482 - 497.

Sheldon, K.M., Kasser, T., Smith, K. & Share, T. (2002). Personal goals and psychological growth: Testing an intervention to enhance goal attainment and personality integration. *Journal of Personality*, 70(1), 5 - 31.

Shin, C., Lee, S. H., Han, K.M., Yoon, H.K. & Han, C. (2019). Comparison of the usefulness of the PHQ-8 and PHQ-9 for screening for major depressive disorder: analysis of psychiatric outpatient data. *Psychiatry Investigation*, 16(4), 300 – 305.

Silber, E. & Tippet, J. (1965). Self-esteem: Clinical assessment and measurement validation. *Psychological Reports*, 16, 1017 - 1071.

Spitzer, R.L., Kroenke, K., Williams, J.B.W. & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, 166, 1092 - 1097.

Torrey, W.C., Mueser, K., McHugo, G.H. & Drake, R.E. (2000). Self-esteem as an outcome measure in studies of vocational rehabilitation for adults with severe mental illness. *Psychiatric Services*, 51, 229 - 233.

Toussaint, A., Hüsing, P., Gumz, A., Wingenfeld, K., Härter, M., Schramm, E. & Löwe, B. (2020). Sensitivity to change and minimal clinically important difference of the 7-item Generalized Anxiety Disorder Questionnaire (GAD-7). *Journal of Affective Disorders*, 265, 395 - 401.

Twohig, M.P., Hayes, S.C., Plumb, J.C., Pruitt, L.D., Collins, A.B., Hazlett-Stevens, H. & Woidneck, M.R. (2010). A randomized clinical trial of acceptance and commitment therapy versus progressive relaxation training for obsessive-compulsive disorder. *Journal of Consulting and Clinical Psychology*, 78(5), 705 - 716.

Twohig, M.P., Shoenberger, D. & Hayes, S.C. (2007). A preliminary investigation of acceptance and commitment therapy as a treatment for marijuana dependence in adults. *Journal of Applied Behavior Analysis*, 40(4), 619 - 632.

Welsh Government (2012). *Psychological Therapies in Wales. Policy Implementation Guidance* <http://wales.gov.uk/docs/dhss/publications/120326psychologicalen.pdf> [Accessed 2<sup>nd</sup> November 2020]

Wicksell, R.K., Ahlqvist, J., Bring, A., Melin, L. & Olsson, G.L. (2008). Can exposure and acceptance strategies improve functioning and life satisfaction in people with chronic pain and whiplash-associated disorders (WAD)? A randomized controlled trial. *Cognitive Behaviour Therapy*, 37(3), 169 - 182.

Wicksell, R.K., Renöfält, J., Olsson, G.L., Bond, F.W. & Melin, L. (2008). Avoidance and cognitive fusion—central components in pain related disability? Development and preliminary validation of the Psychological Inflexibility in Pain Scale (PIPS). *European Journal of Pain*, 12(4), 491 - 500.

Wise, E.A. (2004). Methods for analysing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendations for future directions. *Journal of Personality Assessment*, 82, 50 - 59.

Zettle, R.D., Rains, J.C. & Hayes, S.C. (2011). Processes of change in acceptance and commitment therapy and cognitive therapy for depression: A mediation reanalysis of Zettle and Rains. *Behavior Modification*, 35(3), 265 - 283.