**Work on refining the ACT Knowledge Questionnaire (16 item)**

**Summary**

Using the data from Richards et al. (2011) (Joe Oliver, South London), combined with an unpublished data set from Mindfulness Ltd (Henry Whitfield) and the Luoma & Plumb Vilardarga paper of 2011, I have worked on refining the measure using psychometric analyses.

I have taken two parallel approaches; one using classical test theory (Exploratory Factor Analysis, Item Total Correlation, Cronbach’s Alpha, convergent and divergent correlations) the other using Item Response Theory.

**The data:**

In total we have 212 participants pre training completing the AKQ. We have some other shared measures with which to do convergent correlations. There are other measures that are specific to each data set with an n of 73 and 121 and 20 each. Within each sample here are small amounts of missing data. These have sometimes led to different values of n for different analyses, or when it is a single item I have prorated using the other scores for that individual (case mean substitution) or replaced with the sample mean (sample mean substitution). The amount of data imputed in this way was very low (less than n = 5).

There is some uncertainty and a wide variation in samples of the length of time to follow up: unknown for Mindfulness Ltd, immediately post and one year for the South London Data and immediately post and three months for the US data.

|  |  |  |  |
| --- | --- | --- | --- |
| Sample | N at pre training | n at post training and time | n at follow up and time |
| Mindfulness ltd | 121 | 28  (unknown but likely a range?) | - |
| South London | 73 | 73 (immediate) | article describes n= 24, @ 12 months but I don’t have that data |
| Portland / Reno | 20 | 20 - immediate | 20 – three months |
| Combined | 211 | 121 - unknown | - |

1. Mindfulness Ltd Data

n = 121 collected between 2009 and 2011, a variety of different professionals prior to training.

Demographics: gender, age, profession, therapeutic orientation, use of ACT, current supervision.

Other indices of ACT knowledge: books read, self-rated knowledge,

Specific books read

Number of ACT books read

AAQ9, GHQ12,

Previous training:

|  |  |
| --- | --- |
| 0 | None |
| 1 | A half day or a full day of training |
| 2 | A two day introductory experiential workshop |
| 3 | More than a two day introductory workshop |

Self-rated knowledge

“How would you rate your knowledge of ACT?”

|  |  |
| --- | --- |
| 0 | I have very little or no knowledge about this topic |
| 1 | I have some familiarity with this topic |
| 2 | I have a fairly good grasp of the main principles and strategies |
| 3 | I consider myself competent to apply ACT to a variety of clinical situations |
| 4 | I would consider myself a knowledge expert on this topic |

1. South London Data

n = 73 pre and post training and 24 at one year follow up, collected around 2010,

Training = a one-day workshop

AKQ, AAQ-II,

Self-rated knowledge on a 0 – 10 scale anchored from none to extensive

Number of ACT books read

Number of ACT articles read

Previous training – yes or no and a description of the workshop that we can recode into the same variable as the Mindfulness Ltd and Portland / Reno data

Ratings of likelihood to continue to read about or seek further ACT training

Ratings of experiential willingness to do exercises, share feelings etc in the workshop

1. Portland / Reno Data

Demographics:

n = 20 pre and post a two-day training, with ten randomised to receive six telephone consultations over three months

AKQ, AAQ-II, Satisfaction with training / consultation

Same self-rated knowledge and previous training items as the Mindfulness Ltd data

ACT books read

**Analyses:**

**Item Response Theory: Description of IRT if you are unfamiliar with it:**

IRT analyses the pattern of individual and collective responses to the individual and set of items. It calculates a score for each person’s ‘ability’ on the set of items (basically like a factor analysis of the items and giving each person a factor score for how many they get right. It then also gives a score for each item based on the people passing or failing each item and the ‘ability level’ of those passing and failing. This basically leads to 2 parameters: the item difficulty and the item discrimination. The item difficulty is an estimate of the items location along the ability scale (where does it work best), and the item discrimination is how well it discriminates between people. A third parameter can be calculated giving the likelihood of correct responding by guessing for each item.

These get plotted onto graphs of item characteristic curves and so you can judge how well each item discriminates at different levels of population ability. You are looking for a classic S shaped curve, ideally where an item gives a person of average ability a 50% chance of passing, and that the item gets more likely to be passed with higher ability responders and less likely to be passed at lower ability responders.

You can calculate a test information coefficient, which gives a metric of where you get the most information about your respondents, for example, does the scale give most information about people who are average, above average or below average in ability.

Finally, you can calculate a test characteristic curve that shows the test score associated with the underlying ability scale, which allows you to see where a particular person’s raw score sits relative to the rest of the sample. It also allows you to see how well the whole test works across the range of ability.

**The AKQ Data under IRT**

These are the values for the 3 parameter model (guessing function)

Coefficients:

Guessing Dffclt Dscrmn

AKQ1 0.000 0.724 0.550

AKQ2 0.000 0.405 0.582

AKQ3 0.217 1.819 2.793

AKQ4 0.246 1.336 13.305

AKQ5 0.001 0.426 0.495

AKQ6 0.033 -0.528 1.111

AKQ7 0.317 0.132 1.116

AKQ8 0.000 0.051 1.978

AKQ9 0.000 -0.070 1.320

AKQ10 0.000 -0.381 1.761

AKQ11 0.000 -0.431 1.107

AKQ12 0.000 0.067 0.919

AKQ13 0.000 0.778 1.026

AKQ14 0.000 0.324 1.356

AKQ15 0.000 -0.067 1.375

AKQ16 0.000 0.821 0.657

The above estimates tell us that the probability of correctly guessing the correct answer for items 3, 4, and 7 are relatively high (.22, .25 & .32 – i.e. between 22 and 32 % chance of a correct guess) and so these items increase the likelihood that people will score correct by chance and should be removed as a first step.

Given the sample size, more complex models such as the 3PL model are not advised to be used for text construction, only as diagnostic of high guessing items. Therefore, after this initial run, we will use the 2PL model.

Combined n = 211 sample IRT analysis using 2 Parameter model

Dffclt Dscrmn

AKQ1 0.77662653 0.5136174

AKQ2 0.42647297 0.5564490

AKQ3 4.68494205 0.2251006

AKQ4 2.19522910 0.3700949

AKQ5 0.46543945 0.4533675

AKQ6 -0.59750205 1.0731037

AKQ7 -0.87558431 0.7254299

AKQ8 0.05114625 1.9834056

AKQ9 -0.06605084 1.3314118

AKQ10 -0.37946345 1.7688825

AKQ11 -0.43043142 1.0948648

AKQ12 0.07238593 0.9189042

AKQ13 0.80503711 0.9888640

AKQ14 0.33024523 1.3435188

AKQ15 -0.06620600 1.3286054

AKQ16 0.79356605 0.6924356



An online tutorial by Erin Buchanan gives a cut off figure of above 1 for the discrimination parameter to retain an item. I can’t find a citation for that cut off, in fact I can’t find any guidance on suggested cut offs for retaining or discarding items) but the table above (which is from a good source) would suggest a cut off of greater than one would be on the upper side of the ‘moderate discrimination’ index. This basically means the item does quite a good job of separating people of different ability. Clearly given there is a lack of consensus, we could go higher and remove more items or lower and retain more items.

According to the cut off of 1 however, it suggests we should discard items (1, 2, 3, 4,5, 7, 12, 13, 16).

Looking at the graphs for these items (below), you can see that they don’t discriminate well. The graph on the far right line 1, below shows The Item Characteristic curve for Item 3. For item 3, people of about average ability (0 on the x axis) have about a 22% chance of getting this item correct. People who are 4 standard deviations lower on the ability of ‘ACT Knowledge’ have about a 18% chance of getting it right, and people who are at the very top end of ACT Knowledge Ability have only a 40% chance of getting it right. So this tells us that the item doesn’t discriminate well between people who have a lot of ACT knowledge and those who don’t. Item 12 and 13 are marginal, they have the right shape, but the slope (the discrimination) is not so steep. This is borne out by their discrimination coefficients of .92 and .99. They marginally missed the cut off of 1.

**n = 211, 16 item 2PL IRT, Individual Item Characteristic Curves**

**Items that should be discarded due to discrimination criteria less than one.**

Combined%20211%20sample/211sampleIRTanalysis/Original%2016%20item%20scale/ICC_4_16.pdf

Combined%20211%20sample/211sampleIRTanalysis/Original%2016%20item%20scale/ICC_5_16.pdfCombined%20211%20sample/211sampleIRTanalysis/Original%2016%20item%20scale/ICC_7_16.pdf

/Users/dgilland/OneDrive - University of Edinburgh/Documents OD/Own Research/ACT Training Research/Knowledge Quiz/Combined 211 sample/211sampleIRTanalysis/Original 16 item scale/ICC_12_16.pdf/Users/dgilland/OneDrive - University of Edinburgh/Documents OD/Own Research/ACT Training Research/Knowledge Quiz/Combined 211 sample/211sampleIRTanalysis/Original 16 item scale/ICC_13_16.pdf/Users/dgilland/OneDrive - University of Edinburgh/Documents OD/Own Research/ACT Training Research/Knowledge Quiz/Combined 211 sample/211sampleIRTanalysis/Original 16 item scale/ICC_16_16.pdf

Items remaining in are: 6, 8, 9, 10, 11, 14, 15 – AKQ7item

For the 7 item IRT (2PL) the parameter estimates are:

Dffclt Dscrmn

AKQ6 -0.57349634 1.1789263

AKQ8 0.02832221 1.5930582

AKQ9 -0.07459219 1.6094817

AKQ10 -0.42741142 1.5848214

AKQ11 -0.51419627 0.8924277

AKQ14 0.29599576 1.5155919

AKQ15 -0.07491592 1.5969863

Suggesting these are all working in the middle of the ability range, provide good discrimination within that middle range, less so at the ends.

The ICC plots look like this:

../ICC6_7.pdf../ICC8_7.pdf../ICC9_7.pdf

../ICC10_7.pdf../ICC10_7.pdf../ICC11_7.pdf

../ICC14_7.pdf../ICC15_7.pdf

../AKQ7%20N%20211%20ICC%20plot%20all.pdf../Test%20Information%20Function%207%20items.pdf

The Item Characteristic Curves all look as they should, combining them on one graph (the coloured graph above) shows that the items operate effectively in discriminating between people between -2SD and +2 SD of ability, i.e. the middle range. This is corroborated by the Test Information Function, which shows we are getting the most information in the middle range of the ability scale, and less at both high and low values of ACT knowledge. This tell us that this scale will not work well for people of either very high or very low ACT knowledge. Those with very low ACT knowledge will get a very low score, those with high ACT knowledge should get a near to ceiling effect (i.e. almost all correct). The fact that there is some distance between Item 6 and Item 14 on the X axis tells us that the items aren’t all in exactly the same range of difficulty and there is some spread, which is good. These items are basically neither too hard nor too easy, for people of average ACT knowledge, prior to training.

**Experimenting with a 9 item scale**

Being a little less rigid about the discrimination cut off of 1, and retaining items 12 and 13, we get an IRT coefficient estimate as follows:

Dffclt Dscrmn

AKQ6 -0.58962726 1.1103000

AKQ8 0.04054478 1.8260439

AKQ9 -0.06921264 1.4288934

AKQ10 -0.40194963 1.6835806

AKQ11 -0.44849456 1.0560649

AKQ12 0.07137615 0.8388165

AKQ13 0.79794818 0.9976894

AKQ14 0.31479488 1.4281404

AKQ15 -0.07208595 1.3643855

Items to be retained:

../9%20item%20scale/ICC_9_6.pdf../9%20item%20scale/ICC_9_8.pdf../9%20item%20scale/ICC_9_9.pdf../9%20item%20scale/ICC_9_10.pdf

../9%20item%20scale/ICC_9_11.pdf../9%20item%20scale/ICC_9_12.pdf../9%20item%20scale/ICC_9_13.pdf../9%20item%20scale/ICC_9_14.pdf

../9%20item%20scale/ICC_9_15.pdf

../9%20item%20scale/ICC_9_211_all.pdf../9%20item%20scale/Test%20Information%20Function.pdf

Again, all items look good, so does the combined (colour) plot of the items. In fact, retaining item 13 moves the location difficulty over to the right, meaning that we are getting better discrimination of people with higher ability (i.e. item 13 is a harder item that item 14, it operates at an ability level of .8 of a SD above the average ACT knowledge, whereas item 14 (the most difficult item in the 7 item) only operates at a level of .3 above the average ability). Arguably item 12 is less useful, but it doesn’t appear to be very poor. Its slightly less discriminatory between ability levels, (the slope is less steep). But it is also not 100% passed by people at very high ability and it also isn’t 100% failed by people of very low ability. These would be arguments to retain it.

Test Characteristic Curve

The TCC or Test Response Function (TRF) indicates the total score that is equivalent to the underlying ability score. For the 16, 9 and 7 item scale they looks like this:

16 item 9 item 7 item

../9%20item%20scale/9item%20Test%20reponse%20function.pdf../Test%20response%20function%207%20items.pdf

These figures show that the total score increases smoothly with ability for the 9 and 7 item versions, compared to the 16 item version which has a near vertical slope at scores of 8 to 10. This means that at scores of 8 – 10 on the 16 item scale we aren’t really discriminating between abilities very well. In addition, the maximum possible score on the 16 item version is 16, but in our sample the maximum is just less than 12, suggesting people are not able to score along the full range of the scale (basically there are items that are so difficult that no one can get them right). In contrast, the 7 and 9 item versions score across the full range, with only people of higher ability getting all 9 right (above 2SD higher than average) and people of lower than 2 SD below average scoring zero or 1. Both the 9 and 7 versions show smoothly increasing functions in which scores increase evenly with ability level. The 7 item has a marginally better discrimination profile in that it has a steeper slope in the mid-section, but I don’t think it is of significance in influencing whether we would choose the 7 over the 9.

**Summary of IRT analysis**

We removed items that were likely to be guessed correctly, of the remaining items, we can retain either a 7 or 9 item scale, which is psychometrically better than the 16 item scale. These scales both operate best in the mid-range of ability for people prior to training, they will discriminate well in this region, but less well in discriminating between people who have a very high or very low level of ACT knowledge.

**Classical Test Theory**

The classical test theory represents a slight challenge to the contextual behavioural scientist. It uses language that is ontological such as latent trait, and infers that behaviours (responses to the questionnaire) are driven by the latent trait. As such it can feel mechanistic. I have resolved this tension by remembering that when we ask someone to complete a measure we are basically measuring their behaviour of circling numbers on a page. We make the inference / assumption that that behaviour is associated with a particular functional behaviour, plus error. The behaviour in question will be specified by contextual cues in the measure itself and other features of the testing environment. So for example when someone responds to the CFQ ‘I get very entangled in my thoughts” the cue (the language of the item) leads the person to evaluate how true that is for them, they decide on a number and circle it. The assumption I make is that that behaviour of which number to circle is at least partially influenced by the persons learning history, their awareness, and the actual probability of them responding in a fused way across a variety of circumstances (plus measurement error).

Classical test theory relies on patterns of covariance between such behaviours to say – “when people circle item X in this way, they are also more likely to circle item Y in this way”. Modern computers are able to calculate these degrees of association, determine if they represent covariance of responding above a chance level and then put items together that seem to go together. They specify that the shared variance can be pulled together and called a ‘factor’ (latent variable), and they can specify how much of the responding to each item is influenced by the factor, and by error. This pattern of shared associations and item loadings is Factor Analysis. It has a longer history than IRT, but it has been criticised as ignoring the item variance to concentrate more on the scale variance (hence the development of IRT). Often the solutions that each approach give rise to are similar (sometimes identical). There is a logical sequence of steps in a classical test theory analysis. The results of these steps for the AKQ are described below:

*Apriori theorising:*

It’s worth stating the assumption prior to analysis that we are expecting the AKQ to comprise of one broad ability ‘ACT knowledge’, even if that might be made up of knowledge across a number of domains. So we begin by looking at the item-total correlations and removing any items that correlate with the total scale at <.3 (Nunally, 1978)

|  |  |  |
| --- | --- | --- |
|  | Corrected Item-Total Correlation | Cronbach's Alpha if Item Deleted |
| AKQ1 | .235 | .707 |
| AKQ2 | .225 | .709 |
| AKQ3 | .099 | .720 |
| AKQ4 | .159 | .715 |
| AKQ5 | .175 | .714 |
| AKQ6 | .361 | .694 |
| AKQ7 | .323 | .698 |
| AKQ8 | .490 | .679 |
| AKQ9 | .377 | .692 |
| AKQ10 | .473 | .681 |
| AKQ11 | .372 | .692 |
| AKQ12 | .337 | .696 |
| AKQ13 | .305 | .700 |
| AKQ14 | .362 | .694 |
| AKQ15 | .402 | .689 |
| AKQ16 | .212 | .710 |

Based on Nunnally’s recommendation we would omit items 1, 2, 3, 4, 5, and 16

*Exploratory Factor Analysis*

Factor analysis involves iterative stages. First we need to determine if the data has enough covariance to suggest it is factorable, this is based on two tests – The Kaiser Meyer Olkin test and the Bartlett’s test. KMO should be >.05 and The Bartlett’s test should be significant:

|  |  |  |
| --- | --- | --- |
| **KMO and Bartlett's Test** | | |
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | .783 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 290.550 |
| df | 45 |
| Sig. | .000 |

Next we need to determine how many factors we should expect to extract. To do this we use a technique called parallel analysis. In a Parallel analysis, the computer generates a random data set with the same number of variables and participants that we have, and subjects that random data to a factor analysis. It determines how many factors exist within that data set just by chance and how strong the factors are. That gives us a figure above which our actual factors in the actual data need to be to be better than random.

*Parallel analysis*

With 10 items, and 211 participants, the critical eigenvalue for a factor to be ‘real’ is .96. According to our data there are 3 factors with eigenvalues larger than .96

A second method that is also used is Velicer’s Minimum Averaged Partial Test (MAP) test. According to the MAP test we should extract 1 factor.

If we try to extract three factors, based on the parallel analysis, the solution does not converge within 500 iterations (which is very unusual and suggests cross loading, lack of clear factors.

|  |  |  |  |
| --- | --- | --- | --- |
| **Factor Matrixa** | | | |
|  | Factor | | |
| 1 | 2 | 3 |
| AKQ6 | .432 | .136 | -.164 |
| AKQ7 | .272 | -.006 | .025 |
| AKQ8 | .590 | .069 | .316 |
| AKQ9 | .497 | .099 | -.020 |
| AKQ10 | .540 | .022 | .155 |
| AKQ11 | .604 | -.704 | -.217 |
| AKQ12 | .364 | -.131 | .179 |
| AKQ13 | .371 | -.100 | .165 |
| AKQ14 | .481 | .336 | .055 |
| AKQ15 | .559 | .338 | -.390 |
| Extraction Method: Principal Axis Factoring. | | | |
| a. Attempted to extract 3 factors. More than 100 iterations required. (Convergence=.002). Extraction was terminated. | | | |

In fact, the output above does suggest a number of cross loadings between the three factors, neither factor two or three really looks strong, factor loadings are poor and stronger on the first factor. I did try using an oblique rotation to improve factor structure, where you allow the factors to correlate, but it did not lead to a better solution. Based on this, we should go with the MAP test and our apriori hypothesis that one factor underlies the data.

|  |  |
| --- | --- |
| **Factor Matrixa** | |
|  | Factor |
| 1 |
| AKQ6 | .439 |
| AKQ7 | .280 |
| AKQ8 | .594 |
| AKQ9 | .517 |
| AKQ10 | .560 |
| AKQ11 | .438 |
| AKQ12 | .367 |
| AKQ13 | .374 |
| AKQ14 | .489 |
| AKQ15 | .506 |
| Extraction Method: Principal Axis Factoring. | |
| a. 1 factors extracted. 5 iterations required. | |

This solution only accounts for 29.2% of the variance of the scale which is not a large amount (but not trivial).

Item 7 is a poor loader (again Nunally, 1978 recommends removing items with loadings <.3. Removing item 7 increases variance explained to 32%.

|  |  |
| --- | --- |
| **Factor Matrixa** | |
|  | Factor |
| 1 |
| AKQ6 | .443 |
| AKQ8 | .583 |
| AKQ9 | .515 |
| AKQ10 | .563 |
| AKQ11 | .431 |
| AKQ12 | .364 |
| AKQ13 | .387 |
| AKQ14 | .496 |
| AKQ15 | .506 |
| Extraction Method: Principal Axis Factoring. | |
| a. 1 factors extracted. 5 iterations required. | |

Interesting that this is the same set of items identified in the IRT analysis 9 item version.

*Alpha analysis*

|  |  |  |
| --- | --- | --- |
| Sample | Scale | Alpha |
| Mindfulness Ltd | 9 item | .67 |
|  | 7 item | .66 |
| South London | 9 item | .73 |
|  | 7 item | .75 |
| Portland / Reno | 9 item | .72 |
|  | 7 item | .66 |
| Combined sample | 9 item | .73 |
|  | 7 item | .71 |

These figures are acceptable. The 9 item edges it slightly and it’s better to report at the combined sample level, but if we were being totally transparent we could provide a range of .67 to .73, with a combined alpha of .73.

*Sensitivity to training*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sample | Scale | Mean pre (SD) | Mean post (SD) | Significance | effect size *d* |
| Mindfulness Ltd  n = 28 | 16 | 9.1 (3.1) | 10.5 (2.4) | t = 3.0 (27)  *p* =.006 | .63 |
|  | 9 | 5.9 (1.9) | 6.7 (1.6) | t = 2.6 (27)  *p* = .015 | .47 |
|  | 7 | 4.8 (1.8) | 5.5 (1.5) | t = 2.3  p = .032 | .44 |
| South London  n = 73 | 16 | 5.8 (3.4) | 8.6 (2.5) | t = 7.5 (72)  *p* <.001 | .92 |
|  | 9 | 3.5 (2.4) | 5.1 (1.6) | t = 6.4 (72)  *p*<.001 | .80 |
|  | 7 | 3.1 (2.1) | 4.6 (1.5) | t = 6.3 (72)  *p*<.001 | .79 |
| Portland / Reno data  n = 20\* | 16 | 9.8 (2.6) | 11.1 (2.0) | t = 2.2  p = .041 (19) | .53 |
|  | 9 | 5.6 (2.4) | 6.6 (1.4) | t = 2.1  p = .047(19) | .52 |
|  | 7 | 4.6 (1.9) | 5.4 (1.0) | t = 1.9  p = .07(19) | .49 |
| Combined  n=119 | 16 | 7.2 (3.7) | 9.4 (2.6) | t = 8.1  p <.001 (118) | .79 |
|  | 9 | 4.4 (2.5) | 5.7 (1.7) | t = 6.9  p<.001 (118) | .69 |
|  | 7 | 3.7 (2.2) | 4.9 (1.5) | t = 6.7  p<.001 (118) | .66 |

\*Sample mean substitution used for the total score for two cases, one with missing data pre and one with missing data post.

*Convergent Validity*

Using the variables that are shared we can combine the samples and calculate the following correlations:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Previous ACT Training | Self-rated ACT Knowledge | Number of books read | Acceptance and Action Questionnaire |
| AKQ16 | .19\*\* | .32\*\*\* | .49\*\*\* | -.10 *ns* |
| AKQ9 | .17\* | .30\*\*\* | .50\*\*\* | -.05 *ns* |
| AKQ7 | .20\*\* | .37\*\*\* | .49\*\*\* | -.02 *ns* |

n = 212 \**p*<.05 \*\**p*<.01 \*\*\**p*<.001

That it doesn’t correlate with AAQ is good evidence of divergent validity (i.e. we would not expect it to be correlated with that and it is not.)

We can also demonstrate other correlations by sample:

*Mindfulness Ltd Data*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | AKQ16 | AKQ9 | AKQ7 |
| Awareness of ACT research | *r =* | .11 | .19 | .18 |
| n=109 | *p =* | .27 | .05 | .06 |
| % of practice ACT | *r =* | .34 | .35 | .33 |
| n=119 | *p =* | <.001 | <.001 | <.001 |
| % of practice CBT | *r =* | .33 | .23 | .28 |
| n=119 | *p =* | <.001 | .011 | <.002 |
| # clients on caseload using ACT | *r =* | .28 | .23 | .22 |
| n=119 | *p =* | .002 | .013 | .015 |
| # clients used ACT in the last month | *r =* | .23 | .24 | .22 |
| n=119 | *p =* | .01 | .01 | .02 |
| % clients using some ACT techniques | *r =* | .37 | .34 | .34 |
| n=119 | *p =* | <.001 | <.001 | <.001 |
| Total hours ACT delivered (estimate) | *r =* | .34 | .35 | .34 |
| n = 115 | *p =* | <.001 | <.001 | <.001 |
| Number of metaphors have used | *r =* | .46 | .39 | .42 |
| n = 119 | *p =* | <.001 | <.001 | <.001 |

The AKQ **DOES NOT** correlate with:

self-rating of difficulty learning ACT, level of organisational support for ACT, comfort delivering ACT individually or in groups, number of years qualified, % of practice described as behavioural, analytic, gestalt, existential, solution focussed, motivational interviewing, person-centred, family systems, number of clients previously treated with ACT, number of processes ever used, GHQ12, or frequency of contact with other ACT clinicians.

In addition, the AKQ differs significantly between people who are in contact with other ACT clinicians or not and those that use the ACT list serve versus those that don’t.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Group | *n* = | Mean (*SD*) | *t*(df) | *p* |
| AKQ16 | No ACT contacts | 58 | 7.1 (3.0) | 4.2(115) | <.001 |
|  | ACT contacts | 59 | 9.3 (2.5) |  |  |
| AKQ9 | No ACT contacts | 58 | 4.3 (2.4) | 3.6(107)\* | <.001 |
|  | ACT contacts | 59 | 5.8 (1.9) |  |  |
| AKQ7 | No ACT contacts | 58 | 3.5 (2.0) | 3.5(115) | =.001 |
|  | ACT contacts | 59 | 4.7 (1.7) |  |  |

*\*the df for these is different because the groups have unequal variance and so we use the “variance not assumed” statistics to correct for that, which have different degrees of freedom.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Group | *n* = | Mean (*SD*) | *t*(df) | *p* |
| AKQ16 | Not on listserve | 95 | 7.9 (2.8) | 2.3(117) | .02 |
|  | Follow listserve | 24 | 9.5 (3.6) |  |  |
| AKQ9 | Not on listserve | 95 | 4.8 (2.1) | 2.3(30)\* | .03 |
|  | Follow listserve | 24 | 6.2 (2.8) |  |  |
| AKQ7 | Not on listserve | 95 | 3.9 (1.8) | 2.1(30) | .047 |
|  | Follow listserve | 24 | 5.0 (2.3) |  |  |

*South London Data*

|  | | AKQ16 | AKQ9 | AKQ7 |
| --- | --- | --- | --- | --- |
| ArticlesRead | Pearson Correlation | .313\*\* | .314\*\* | .356\*\* |
| Sig. (2-tailed) | .008 | .007 | .002 |
| N | 72 | 72 | 72 |
| Interest in Reading more | Pearson Correlation | .065 | .101 | .125 |
| Sig. (2-tailed) | .592 | .402 | .298 |
| N | 71 | 71 | 71 |

*South London Data (continued).*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | AKQ16 | AKQ9 | AKQ7 |
| Interest in getting more training | Pearson Correlation | .189 | .213 | .255\* |
| Sig. (2-tailed) | .114 | .075 | .032 |
| N | 71 | 71 | 71 |
| Interest in using ACT more | Pearson Correlation | .030 | -.053 | .048 |
| Sig. (2-tailed) | .805 | .661 | .688 |
| N | 71 | 71 | 71 |

*Correlations between the different versions*

As you would expect there are very high correlations between the versions, suggesting they are measuring the same thing.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Correlations** | | | | |
|  | | AKQ16TotalPre | AKQ9TotalPre | AKQ7TotalPre |
| AKQ16TotalPre | Pearson Correlation | 1 | .905\*\* | .868\*\* |
| Sig. (2-tailed) |  | .000 | .000 |
| N | 212 | 212 | 212 |
| AKQ9TotalPre | Pearson Correlation | .905\*\* | 1 | .958\*\* |
| Sig. (2-tailed) | .000 |  | .000 |
| N | 212 | 212 | 212 |
| AKQ7TotalPre | Pearson Correlation | .868\*\* | .958\*\* | 1 |
| Sig. (2-tailed) | .000 | .000 |  |
| N | 212 | 212 | 212 |
| \*\*. Correlation is significant at the 0.01 level (2-tailed). | | | | |

**Conclusions**

The scale can be improved, either the 9 item or the 7 item solution is better than the 16 item across a number of parameters. In deciding between the 9 and 7 item version here is how it breaks down

|  |  |
| --- | --- |
| Parameter | Result |
| IRT analysis | Nothing in it. |
| Factor analysis | Variance explained and item loadings favour the 9 item solution |
| Alpha analysis | Favours 9 item solution |
| Sensitivity to training | Favours 9 item solution |
| Convergent validity | Combined sample – not much in it, but slightly favours 7 item |
|  | Mindfulness Ltd data – Out of 16 correlations, 14 are equivalent and 2 slightly favour the 7 (very marginal) |
|  | SLAM data – Of 8 correlations, 2 marginally favour the 7 item, there isn’t much in it, though one of them moves from non-significant to significant: (interest in getting more training is ns in 9 and <.05 in 7.) |

Based on the overall pattern of data I would recommend a 9 item AKQ as the best scale.

*Revised numbering*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 16 item | 6 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 9 item | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Correct | b | d | b | b | d | c | b | d | c |

The AKQ-Revised (AKQ-R):

1. A client tells a story about her life that includes drinking alcohol every day, three failed marriages, moving every 12 months, overeating, and repetitious self-injury. What process is most likely to functionally connect these issues?
2. escape maintained behaviour
3. experiential avoidance
4. relational frames of comparison and time
5. excessive cognitive fusion
6. Which of the following best illustrates a client’s confusion with goals as values?
7. A man wants to be a good employee.
8. An adolescent wants to be more educated.
9. A woman wants to be emotionally available for several people in her life.
10. A woman wants to be married.
11. According to the ACT book, when a therapist says the phrase “If you are not willing to have it, you’ve got it” he is illustrating the concept of
12. defusion.
13. control as the problem.
14. acceptance.
15. values.
16. Which of the following is not an ACT-consistent explanation of “psychopathology”?
17. emotional avoidance.
18. ineffective thinking and behaviour patterns.
19. cognitive fusion.
20. lack of committed action.
21. Ongoing self-awareness is the same as
22. self-as-content.
23. the conceptualized self.
24. the evaluated self.
25. self-as-process.
26. Which of the following is not a statement about contact with the present moment?
27. Thoughts and feelings often present themselves as about the past or future, but they are experienced now.
28. Cultivating awareness of thoughts and emotions as they occur allows us to notice when they get in the way of valued action.
29. You are not your thoughts, memories, or roles.
30. Life is not something to be lived when you have solved your problems, life is going on now.
31. Values are
32. non-verbal qualities of action
33. verbally construed global desired life consequences
34. a decision, not a choice
35. the sum of the goals achieved while on a life path
36. Willingness, as defined by the ACT book, refers to
37. a person’s motivation to try something new or different in their life.
38. a feeling or belief that is helpful for tolerating discomfort.
39. noticing thoughts as verbal constructions.
40. giving up the struggle with emotional discomfort and disturbing thoughts.
41. The purpose of creative hopelessness is:
42. To create a coherent story about why the client’s life is painful.
43. To help a client recognize that his or her life, as it is being lived now, is hopeless.
44. To show that the strategies that the client has used to manage internal experiences are unworkable.
45. To illustrate to the client that they need to find new ways to fix their problems.