Personality and treatment response to electroacupuncture. A new measure of mood change and further analysis of questionnaire response styles

© David Mayor[1*] and Tony Steffert[2]

CONTENTS

[1] School of Health and Social Work, University of Hertfordshire.
* Corresponding author: davidmayor@welwynacupuncture.co.uk.
[2] Music Computing Lab, Open University

2. Secondary self-report measures

2. 1. Response style (RS)
2.2. Responsiveness
2.3. Shannon entropy (SE)
2.3.1. Standard deviation (SD)

Hypotheses tested

Hypothesis 1. The multiple mood change Likert scale (MMCL) shows acceptable validity and internal consistency

1.1. Internal consistency – Cronbach's alpha
1.2. Convergent and discriminant validity

Hypothesis 2. The MMCL is less sensitive to change in mood than pre-to-post differences in NRS-M or $BRUMS_{24}$ scores

Hypothesis 3. Some respondents may consistently report larger or smaller mood changes than the majority

3.1. Skewness of response data
3.2. Test discrimination **delta$_G$**

Hypothesis 4. Mood changes found, and their magnitude and direction, may be affected by personality type or attitude

Hypothesis 5. Benefit from prior acupuncture as assessed using the AHQ may be reflected in mood changes found during the teaching session

Hypothesis 6. Benefit from prior acupuncture may be reflected in EA stimulation parameters described as most or least pleasant and/or intense

Hypothesis 7. Which EA stimulation parameters are described as most or least pleasant and/or intense may be affected by personality type or attitude

Hypothesis 8. MMCL scores may show an association with which stimulation parameters are found to be most or least pleasant and/or intense

Hypothesis 9. Whether acupuncture is reported as beneficial in the past may depend on personality type or attitude

Hypothesis 10. There may be differences in personality type and/or attitude among training institutions

Hypothesis 11. Missing data may depend on personality type, attitude, or training institution

Hypothesis 12. Other significant interactions between primary measures, between primary and secondary measures, and between secondary measures

Hypothesis 13. Some respondents may show similar patterns of responsiveness, response style or Shannon entropy over a number of the mood scales and/or personality/attitude measures

Hypothesis 14. There may be significant interactions among the secondary measures themselves

13/14.1. Response style

Hypothesis 13.1.1. Individuals may complete all questionnaires in a similar way, as assessed by response style

Hypothesis 13.1.2. Individuals may complete all questionnaires in a similar way, as assessed with Shannon entropy

Background

In our previous research into the effects of electroacupuncture (EA) delivered in a classroom situation [Mayor & Steffert 2016; Steffert & Mayor 2017] we have used both Likert scales (the Brunel Mood Scale, $BRUMS_{24}$) and multiple numerical rating scales for mood (NRS-M) to assess moods before and after supervised treatment by a fellow student, with changes in mood being measured by the difference between 'pre' and 'post' scores. In a direct comparison of the $BRUMS_{24}$ and NRS-M)[Mayor & Steffert 2016], we noted that the NRS-M performed better in terms of test discrimination, skewness of responses and amount of change measured, but that the $BRUMS_{24}$ appeared more sensitive to change at a group level, although effect sizes for both mood measures were small.

We also investigated the effects on mood changes of some personality measures (such as emotional intelligence and Jung typology) and the reported helpfulness of prior acupuncture. In addition, we explored the use of secondary measures describing *how* these scales were completed [Steffert & Mayor 2017].

In this further study, we pilot a multiple mood *change* Likert scale (MMCL) instead of a multiple mood NRS, consider a different and extended set of personality measures, continue with the use of secondary measures, and also examine respondents' comparisons of the intensity and pleasantness of the different treatments they received.

Objectives

1. To apply several different primary measures of personality type/attitude and continue our earlier explorations of the helpfulness of prior acupuncture in the teaching situation; 2. To further investigate the secondary measures previously used to describe how the various scales are completed; 3. To pilot the MMCL, examine its validity and internal consistency, and compare its performance with the mood scales we used previously; 4. To investigate which parameters of EA respondents found most or least pleasant or intense; 5. To test a number of hypotheses on the interactions between these various measures.

Our hypotheses were as follows:

1. The multiple mood change Likert scale (MMCL) shows acceptable validity and internal consistency

2. The MMCL is less sensitive to changes in mood than pre-to-post differences in NRS-M or $BRUMS_{24}$ scores

3. Some respondents may consistently report larger or smaller mood changes than the majority

4. Mood changes found, and their magnitude and direction, may be affected by personality type or attitude

5. Benefit from prior acupuncture may be reflected in MMCL mood changes found during the teaching session

6. Benefit from prior acupuncture may be reflected in which EA stimulation parameters are described as most or least pleasant and/or intense

7. Which EA stimulation parameters are described as most or least pleasant and/or intense may also be affected by personality type or attitude

8. MMCL mood changes may show an association with which stimulation parameters are found to be most or least pleasant and/or intense

9. Whether acupuncture is reported as beneficial in the past may depend on personality type or attitude

10. There may be differences in personality type and/or attitude among training institutions

11. Missing data may depend on personality type, attitude, or training institution

12. There may be other significant interactions between primary and secondary measures

13. Some respondents may show similar patterns of responsiveness, response style or Shannon entropy over a number of the mood scales and/or personality/attitude measures

14. There may be significant interactions among the secondary measures themselves.

Methods

Recruitment

Ninety respondents were recruited during six EA teaching sessions in the UK and Denmark, attended as part of the requirement for a qualification in acupuncture practice or as part of a continuing professional development (CPD) programme. The teaching sessions all followed a similar standard pattern, although duration varied between half a day and two days; respondents were encouraged to complete the scales but not obliged to do so. Ethics approval was granted under applications for related studies by the Health and Human Sciences Ethics Committee of the University of Hertfordshire, UK. Permission was also received from the course organisers and respondents themselves.

Measures

1. Primary self-report measures

The following measures were used:

1.1. Acupuncture helpfulness questionnaire (AHQ) [Steffert & Mayor 2017]

1.2. Items from the Technology Attitudes Scale (TAS) [Rosen et al. 2013]

1.3. The Behavioural Approach and Inhibition scales (BIS/BAS) [Carver & White 1994]

1.4. The Body Awareness Questionnaire (BAQ) [Shields et al. 1989]

1.5. The Positive and Negative Affect Schedule, short form (I-PANAS-SF) [Thompson 2007]

1.6. The Jung Typology Test™ (JTT) [Jung & Haynes 1971; Briggs Myers & Myers 1995]

1.7. The Multiple Mood Change Likert scale (MMCL)

1.8. Pleasantness/Intensity of treatment

Of the primary measures, all are based on Likert scales apart from (1.8), which is in multiple choice format. The measures as used can be found in **Appendix B**.

2. Secondary self-report measures

The following measures were used:

2.1. Response style (RS)

2.2. Responsiveness

2.3. Shannon entropy (SE)

A method for calculating SE in Excel can be found in **Appendix A**.

3. Treatment description

In addition, participants were asked to observe and complete a description of the treatment carried out on another person attending the course (not themselves). This data is not discussed further here.

Administration

Participants were requested to complete the self-report measures 1.1 to 1.5 on paper, and measure 1.6 online, *before* attending the training course. Self-report measures 1.7 and 1.8 were completed during the course itself, when attendees had received a brief treatment with EA and/or transcutaneous EA stimulation (TEAS) from a fellow attendee. Measure 3 was completed by another attendee acting as observer.

Some attendees declined the treatment because of known contraindications (e.g. pregnancy or a heart condition), and occasionally because of an aversion to something 'non-traditional' like EA or TEAS, or to electricity itself. Some attendees left sessions early, and thus did not complete all scales. Treatment was supervised but participants were free (within reason) to use their own choice of parameters (frequency, amplitude, mode, pulse and overall stimulation duration) and acupuncture points.

As before, if students had not completed the online Jungian typology self-test (JTT) or other personality/attitude measures before the session (or had completed a version of the Myers-Briggs Type Indicator (MBTI) instead of the JTT), they were asked to do so during a break in the teaching.

Analysis

Response data were entered in numerical form in Excel spreadsheets. If data were missing, the letters 'na' were entered instead. Analysis was conducted using SPSS v 23 and Excel 2010 v 14.0. Questionnaire responses and other data were tested for skewness and kurtosis but none – other than age – were consistently found to be normally distributed. Thus non-parametric methods were used throughout. Other tests for normality (such as the Shapiro-Wilk test) were not used, and data was not transformed.

Results

Respondents

Of the 90 respondents, 76 were women (median age 40, range 21-60) and 13 men (median age 43, range 23-61); two respondents did not state age, and one of these did not state gender. Median age for the whole sample was 40 (thus, for some analyses, age was dichotomised as either < 40 or ≥ 40).

The courses were held at four different training institutions:

- Akupunktur Akademiet (AA), Åbyhøj, Denmark
  (1 CPD course, 16 participants)

- College of Integrated Chinese Medicine (CICM), Reading, UK
  (2 undergraduate courses, 38 students)

- London South Bank University (LSBU), London, UK
  (1 undergraduate course, 8 students)

- Northern College of Acupuncture (NCA), York, UK
  (2 undergraduate courses, 28 students).

Missing data

There was rather more missing response data for respondents from CICM than from the other training institutions, and least from the CPD course. Median age was highest for the CPD course, and lowest for the course held on University premises (**Table 1**).

Table 1. Numbers of participants at the different training institutions (*N*), showing number of 'na' missing data entries (*n*), median age and number of women and men attending.

Training institution	*N* participants	*n* 'na' responses	*n/N*	median age	*N* F; M
AA	16	236	14.8	47	11; 5
CICM	38	1136	29.9	40	32; 5
LSBU	8	153	19.1	28	8; 0
NCA	28	484	17.3	42½	25; 3
Totals	90	2009	22.3	40	76; 13

Because of a mishap with questionnaire printing at CICM, most of the missing data was for the MMCL there, closely followed by the JTT, which respondents were required to complete in their own time (**Table 2**). However, possibly because students at CICM are more routinely exposed to the MBTI, a lower percentage of CICM attendees than NCA attendees failed to complete the JTT: 18.4% (7) for CICM, as against 28.6% (8) for NCA.

Table 2. Missing data by questionnaire, showing number of 'na' missing data entries (*n*), number of whole questionnaires completely missing ('no response') and number of items per questionnaire (*q*).

Questionnaire	*n* 'na' responses	no response	*q*	*n*/*q*
AHQ	58	6	5	11.60
TAS	7	0	8	0.88
BIS/BAS	37	0	24	1.54
BAQ	23	0	18	1.28
I-PANAS-SF	33	4	10	3.30
JTT	1216	19	64	19
MMCL	217	22	10	21.70

Thus results for the MMCL, JTT and AHQ, in particular, should be treated with caution unless confirmed by further study.

There was less missing data in the questionnaires completed by the 46 older than by the 44 younger attendees (13.9 items per individual for the former, 18.5 items per individual for the latter – a nonsignificant difference). There was also less missing data for women (15.32 items per individual) than for men (16.92 items per individual) – again a nonsignificant difference.

Questionnaire responses

1. Primary self-report measures

1.1. Acupuncture Helpfulness Questionnaire (AHQ)

This was used to assess attendees' recollection of the helpfulness of prior acupuncture received.

The possible range of scores for each item in this questionnaire was from 1 ('Not [helpful] at all') to 7 ('Extremely helpful'), the median score being 4.

Eighty-nine respondents stated that they had received acupuncture treatment in the past. Of these, 80 (92%) scored > 4 in response to the question whether they had found acupuncture 'generally helpful'. For the other, more specific questions, responses were more restrained, with around 68-69% of respondents scoring > 4 in response to questions on whether they had found acupuncture helpful for purely Physical/biomedical or Functional/unexplained disorders or for the Lifestyle advice given, but 78% scoring > 4 for the question on mental/emotional difficulties.

There were no significant differences between men and women or between older and younger respondents in AHQ responses, although the greatest difference was for Functional/unexplained disorders, for which women found treatment more helpful than men, although not significantly so (median values 6 and 4.5, respectively, $p = 0.157$, effect size $r = 0.16$).

Analysing the AHQ responses using a hierarchical cluster analysis for a range of clusters between 2 and 4 (with Ward's method and squared Euclidean distances), the clearest result is a 2-cluster solution, with those who found acupuncture particularly helpful for mental/emotional difficulties as a single-item one cluster, and the other three items in the other. This is confirmed by the values of Spearman's *rho* for bivariate correlations between the Physical and Functional items (*rho* = 0.592), with lower values – still significant at the 0.01 level – for correlations between these and the

Lifestyle item. In contrast, correlations of the Physical and Functional items with the Emotional/mental item were only significant at the 0.05 level, whereas the correlation between the latter and the Lifestyle item was again at the 0.01 level. This suggests that although there is considerable overlap in the conditions for which future acupuncturists attend treatment, there are two broad categories – Physical and Functional versus Emotional/mental.

1.2. Items from the Technology Attitudes Scale (TAS)

Mean and standard deviation of scores for positive attitude to technology were 3.28 (0.65), and 3.42 (0.73) for negative attitude to technology. The first of these is significantly less than the expected mean value (3.66) for the *complete* scale found in the literature [Rosen et al. 2013] ($t = -5.536$, $df = 89$, $p < 0.001$, using a one-sample T-test), whereas there is no significant difference between the mean for negative attitude to technology and the normative value found in the literature. This suggests – perhaps not surprisingly – that in this sample of traditional acupuncture students and practitioners, attitude to technology may be less positive than in the general population.

Women had a slightly more positive attitude to technology than men, but not significantly so, and younger respondents a more negative attitude than older respondents, but again this did not reach significance ($p = 0.058$, $r = 0.20$).

1.3. Behavioural Approach and Inhibition scales (BIS/BAS, or BB)

The 24-item BIS/BAS scales were devised to assess individual tendencies to move towards something desired or away from something considered unpleasant [Carver & White 1994].

Using established categories of low, medium and high scorers from the literature [Carver & White 1994], numbers in this study for each category were as shown in **Table 3**.

Table 3. Numbers of low, medium and high BIS/BAS scorers in this study (*N* = 89). Ranges are indicated in square brackets

Subscale	***N* Low scorers [range]**		***N* Medium scorers [range]**		***N* High scorers [range]**	
BAS Reward Responsiveness	18	[5-14]	61	[15-18]	10	[19-20]
BAS Drive	12	[4-8]	65	[9-13]	12	[14-16]
BAS Fun seeking	11	[4-9]	68	[10-14]	10	[15-16]
BIS	20	[7-16]	66	[17-23]	3	[24-28]

Thus only for BIS (behavioural inhibition) was there a significant difference between numbers of low and high scorers ($p<0.001$, using the Binomial test for expected ratio 0.50). Thus this sample of acupuncture students and practitioners may exhibit more rather than less exploratory behaviour.

However, using a 1-sample T-test for this sample, there are significant differences from expected mean values (weighted averages taken from questionnaire results from the different groups investigated in the original publication by Carver and White 1994) (**Table 4**).

Table 4. Differences between BIS/BAS scores here and normative data, using a 1-sample T-test.

Subscale	Average (SD)	Normative data	Difference significant
BAS Reward Responsiveness	16.06 (2.10)	17.63 (2.10)	$p < 0.001$
BAS Drive	11.18 (2.41)	12.05 (2.42)	$p = 0.001$
BAS Fun seeking	11.71 (2.18)	12.44 (2.24)	$p = 0.002$
BIS	18.43 (3.00)	20.01 (3.81)	$p < 0.001$

Thus for BIS and two of the BAS scales, this group showed significantly more 'behavioural activation' than would be expected from prior research, but less for the third BAS scale, Reward responsiveness – perhaps offering an interesting insight into the psychology of acupuncturists.

There were no significant differences between women and men in the BAS and BIS scales in this study, although scores on all scales were marginally higher for the former. Differences were also consistently greater between younger and older respondents, with the former scoring more highly – significantly so for 'Fun seeking' ($p = 0.001$; Mann-Whitney $U = 578.5$, $Z = -3.278$, $r = 0.35$), and almost so for 'Reward Responsiveness' ($p = 0.054$, $r = 0.21$).

1.4. Body Awareness Questionnaire (BAQ)

The BAQ is an 18-item scale designed to assess self-reported attentiveness to normal non-emotive body processes, such as cycles and rhythms, the ability to detect small changes in normal functioning, and the ability to anticipate bodily reactions [Shields et al. 1989].

The minimum and maximum possible scores on the BAQ are 18 and 126 (median 66.5). Most (81) respondents scored more than this median, and only 8 below it. Nonetheless, using a 1-sample T-test, the average BAQ score and its standard deviation was considerably *lower* than those reported in the literature [Shields et al. 1989], both for women ($p < 0.001$) and men ($p = 0.010$), and for younger ($p < 0.001$) and older ($p < 0.001$) respondents (**Table 5**).

Table 5. Differences between BAQ scores and normative data, using a 1-sample T-test.

BAQ	Average (SD)	Normative data	Difference significant
Women	87.2 (15.5)	107 (19.3)	$p < 0.001$
Men	86.0 (15.0)	98.7 (21.3)	$p = 0.010$
Younger	87.5 (12.9)	103.7 (18.75)	$p < 0.001$
Older	86.7 (17.6)	102 (21.85)	$p < 0.001$

This is a somewhat surprising result, given that acupuncturists are, in some ways, trained to be aware of their own and their patients' bodies and responses. It is, of course, possible that the BAQ does not reflect this kind of acupuncturalawareness.

BAQ scores did not differ significantly with age or gender, although they were slightly higher for women than men, and for younger than older respondents.

1.5. Positive and Negative Affect Schedule, short form (I-PANAS-SF)

There are several variants of scales designed to assess positive and negative emotional style. Here we used a short form devised by Thompson [2007].

The minimum and maximum possible for the positive and negative affects in this scale are 5 and 25 (median 15). Most respondents (77) scored higher than the median for positive affect, and most (82) scored lower than the median for negative affect. Average values were quite similar to those reported in the literature [Thompson 2007], but using a one-sample T-test the difference was significant for positive affect (18.40 versus expected mean 19.71; $p < 0.001$), although not for negative affect (11.30 versus expected mean 11.27; $p = 0.910$).

Thus, according to the I-PANAS-SF, respondents showed more positive than negative affect, although they were less positive than expected, and slightly more negative.

1.6. Jung typology test™ (JTT)

One of the most popular personality typologies was devised initially by Carl Jung [Jung & Haynes 1971], and developed commercially by Katharine Cook Briggs and Isabel Briggs Myers [Briggs Myers & Myers 1995]. Here we used the open-access online Jung typology test™ (JTT).

As described in our previous research, there are eight main and 16 combination subtypes in the JTT [Steffert & Mayor 2017]. The numbers of types and subtypes in this sample are shown in **Tables 6** and **7**.

Table 6. Numbers of different JTT types in this sample, and by training institution, showing significance and relative risk for a Binomial test on the total counts (test ratio 0.5).

#	Jungian (JTT) type	code	AA	CICM	LSBU	NCA	**Totals**	**Binom**
1	Introversion (I)	1	5	14	4	13	36	n.s.
2	Extraversion (E)	2	8	17	2	7	34	(RR 2.1)
3	Intuition (N)	1	7	21	**6**	11	45	p=0.022
4	Sensing (S)	2	6	10	0	9	25	(RR 3.6)
5	Feeling (F)	1	**10**	**25**	5	**17**	57	p<0.001
6	Thinking (T)	2	3	6	1	3	13	(RR 8.8)
7	Perceiving (P)	1	4	18	4	13	39	n.s.
8	Judging (J)	2	9	13	2	7	31	(RR 2.5)
	Totals		52	124	24	80	280	

Thus in this sample – in line with findings in our 2017 study – there are significantly more Intuition than Sensing types, and more Feeling than Thinking types. Furthermore, there are more Introverts than Extraverts among those attending the NCA courses, whereas this is not the case for the CICM courses. This was also a nonsignificant finding in our previous study [*ib*.].

Table 7. Numbers of different JTT combination subtypes in this sample, and by training institution.

#	Jungian (JTT) combination type	AA	CICM	LSBU	NCA	**Totals**
1	ISTJ	1	0	0	1	2
2	ISFJ	1	1	0	2	4
3	INFJ	1	3	1	3	8
4	INTJ	0	1	0	0	1
5	ISTP	0	1	0	1	2
6	ISFP	0	2	0	4	6
7	**INFP**	2	**5**	**2**	2	11
8	INTP	0	1	1	0	2
9	ESTP	0	1	0	0	1
10	ESFP	0	1	0	0	1
11	**ENFP**	1	**7**	1	**6**	15
12	ENTP	1	0	0	0	1
13	ESTJ	1	2	0	1	4
14	ESFJ	**3**	2	0	0	5
15	ENFJ	2	4	1	0	7
16	ENTJ	0	0	0	0	0
	Totals	13	31	6	20	70

As in our previous small study of those attending EA courses [*ib*.], there are more Feeling-Perceiving subtypes (whether Introvert or Extravert) than any of the others – a similar result to that found for counsellors [*ib*.].

Adding the results for the two studies together – both using exactly the same methods to gather the JTT data – results in **Table 8**.

Table 8. Total numbers of different JTT types in this *and* the previous sample, and by training institution, showing relative risk (RR) rather than effect size.

#	Jungian (JTT) type	AA	CICM	LSBU	NCA	**Totals**	**Binom**
1	Introversion (I)	5	30	10	27	72	n.s.
2	Extraversion (E)	8	31	8	16	63	(RR 2.3)
3	Sensing (S)	7	30	9	20	66	n.s.
4	Intuition (N)	6	31	9	23	69	(RR 1.9)
5	Feeling (F)	10	45[a]	15[b]	39[a]	109	p<0.001
6	Thinking (T)	3	16	3	4	26	(RR 8.4)
7	Perceiving (P)	4	32	9	25	70	n.s.
8	Judging (J)	9	29	9	18	65	(RR 2.2)
	Totals	52	244	72	172	540	

a. p<0.001; b. p=0.009

1.7. Multiple mood change Likert scale (MMCL)

Results for this scale are analysed in detail below, under 'Hypotheses tested'.

1.8. Pleasantness/Intensity of treatment

Several respondents found that stimulation parameters were not necessarily unique in being experienced as 'most intense', 'least pleasant and so forth. For example, both low frequency (LF) *and* high frequency (HF) were scored as 'most intense' by two respondents, and high frequency *and* dense/disperse stimulation (DD) by another two. When such tied results were split into their component parameters, results were as shown In **Table 9**.

Table 9. Pleasantness and intensity of treatment reported by attendees.

Parameters	Most intense	Least intense	Most pleasant	Least pleasant	Totals
Low frequency (LF)	18	21	23	17	79
Midrange (MF)	4	12	14	8	38
High frequency (HF)	25	12	16	27	80
Dense/disperse (DD)	11	12	20	9	52
Totals	58	57	73	61	249

Thus no particular stimulation parameters stand out for either intensity or pleasantness, although HF stimulation was more often reported as most intense and least pleasant, LF as most pleasant and least intense.

2. Secondary self-report measures

2. 1. Response style (RS)

Response style is a respondent's tendency to respond to questionnaires in certain ways regardless of their content [Van Vaerenbergh & Thomas 2013]. RS has been researched since the 1960s, and in the last 14 years the amount of such research published each year has increased exponentially (PubMed search for "response style") (**Figure 1**).

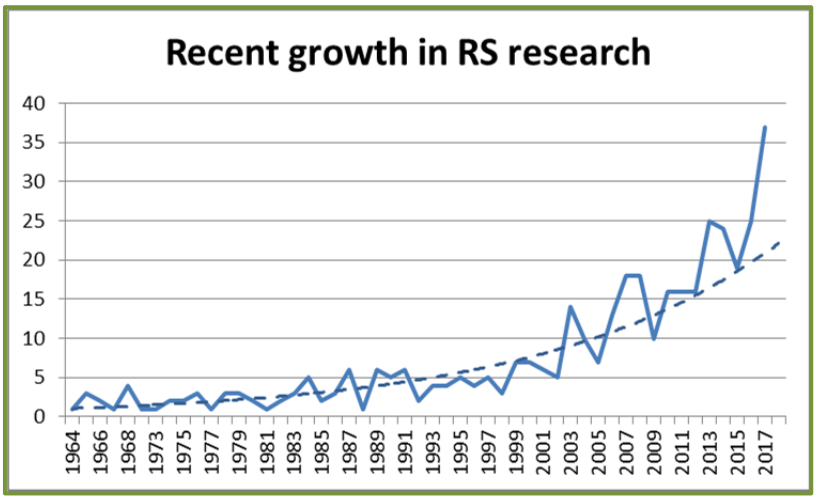

Fig 1. Annual rates of RS publication (Source: PubMed search 24 Jan 2018).

In our previous study [Steffert & Mayor 2017], we used two main RS measures, Extreme Response Style (ERS), or a tendency to score towards the extremes of a scale, and Midpoint Response Style (MRS), or a tendency to score towards the middle of a scale. Scoring methods for the questionnaires used in the present study, similar to those explained there, are shown in **Table 10**, which also shows the median MRS and ERS counts for each questionnaire and all questionnaires taken together.

Table 10. Scoring MRS and ERS for the measures in this study.

Questionnaire	Range of scores	MRS	ERS (low; high)	Median MRS	Median ERS
AHQ	1 to 7	3, 4 or 5	1 or 2; 6 or 7	1	3
TAS	1 to 5	3	1; 5	2	1
BIS/BAS	1 to 4	2 or 3	1; 4	16	7
BAQ	1 to 7	3, 4 or 5	1 or 2; 6 or 7	7	10
I-PANAS-SF	1 to 5	3	1; 5	3	1
JTT	Scored online	3	1; 5		
MMCL	-2 to +2	0	-2; +2	4	1
ALL scales together				35	23.5

Across all measures (AHQ, TAS, BIS/BAS, BAQ, I-PANAS-SF and MMCL), MRS and ERS counts for the different respondents varied between 2 and 52, and 2 and 44, respectively, with 3051 MRS scores and 2207 ERS scores in total.

For TAS, BIS/BAS, I-PANAS-SF and MMCL, there were more MRS than ERS scores, and for AHQ and BAQ, more ERS scores.

There was no significant difference in either MRS or ERS for older versus younger respondents, and no consistent but nonsignificant difference in median MRS or ERS values were found across all scales. However, there were negative correlations between age and ERS for the TAS and BIS/BAS scales (Spearman's *rho* = -0.234 and -0.221, respectively, significant at the 0.05 level), suggesting that for these two measures at least, older respondents may use a less Extreme response style.

In contrast, there were some significant differences by gender. For two scales (AHQ and BIS/BAS) and for all scales taken together, ERS was significantly greater for women than for men (effect sizes 0.21 to 0.27), and for BIS/BAS MRS was significantly greater for men ($p = 0.014$, effect size 0.26).

For all the scales except TAS and MMCL, median MRS was higher for men than for women, and median ERS higher for women than for men. Average ERS for TAS (but not MMCL) was also higher for women than for men.

It thus seems possible that, with a larger sample, this pattern of women using more of an ERS and men more a MRS may be found across several questionnaires. There is lack of agreement on this in the literature about different scales [Van Vaerenbergh & Thomas 2013].

Respondents who scored with a consistent MRS or ERS across three or more questionnaires – i.e. in the upper quartile of the number of MRS or ERS scores for those questionnaires – were as shown in **Table 11.**

Table 11. Respondents in the upper quartile of MRS or ERS counts for the questionnaires used in this study.

Respondent (by number)	MRS [gender, age]	ERS [gender, age]
1	0	1 [F, 45]
2	0	1 [F, 40]
4	0	1 [M, 26]
15	1 [M, 43]	0
18	0	1 [F, 60]
23	1 [F, 52]	0
25	1 [M, 47]	0
37	0	1 [F, 29]
41	0	1 [F, 37]
47	0	1 [F, 25]
49	0	1 [F, 50]
51	1 [M, 61]	0
52	0	1 [F, 55]
59	0	1 [F, 33]
60	1 [F, 39]	0
61	1 [F, 37]	0
63	1 [F, 46]	0
67	0	1 [F, 28]
68	0	1 [F, 21]
71	0	1 [F, 27]
76[a]	1 [F, 46]	0
81	1 [F, 34]	0
82	1 [F, 31]	0
84	1 [M, 38]	0
Totals [*N* F, M; median age]	11 [7 F, 4 M; 43]	13 [12 F, 1 M; 33]

a. Scored with MRS on four questionnaires. The remainder scored with MRS or ERS only on three.

The gender balance for these consistent MRS and ERS scorers is in keeping with the finding on gender above; the difference in median ages is also of interest, suggesting that with a larger sample, age might after all have some impact on RS.

2.2. Responsiveness

Judging as 'responsive' those who responded with a MMCL score of ±2 in a 'good' or 'healthy' direction (e.g., -2 for Anxious or Tired, +2 for Alert or Relaxed) on five or more items, **Table 12** shows those respondents with high responsiveness ('Interpretation 1'). However, because the items Alert and Tired could be construed respectively as 'bad' (hyperalertness) or 'good' (sleepy and comfortable) in some contexts, responsiveness was also calculated with the valence for these two items reversed ('Interpretation 2').

Table 12. Responsiveness of individual attendees. In **bold** and underlined, those individuals who appear both as 'good responders' and 'marginally poor responders'.

Respondent (by number) (5 or more ±2 scores)	**Responsiveness (Interpretation 1)**	**Responsiveness (Interpretation 2)**	**Non-responsive (8 or more zero scores)**	**Poorly responsive (2 or more 'bad' scores)**
			2 [F, 40]	*Interpretation 1*
8 [F, 46]	1	1	4 [M, 26]	**8** [F, 46]
11 [M, 25]	0	1	12 [M, 25]	**37** [F, 29]
14 [F, 47]	1	1	19 [F, 44]	46 [F, 33]
18 [F, 60]	1	0	21 [F, 21]	48 [F, 56]
36 [F, 52]	1	0	25 [M, 47]	54 [M, 25]
37 [F, 29]	0	1	34 [F, 49]	
43 [F, 48]	1	0	57 [F, 53]	*Interpretation 2*
45 [M, 48]	1	0	60 [F, 39]	45 [M, 48]
55 [M, 23]	1	1	82 [F, 31]	48 [F, 56]
75 [F, 32]	1	0		

There is no striking difference in the numbers of men or women who are responsive or non-responsive, and the median ages of responders and non-responders are not dissimilar (46.5 vs 39.5 years).

A zero score could be taken to indicate non-responsiveness, either of the respondents or of the MMCL itself in this context. Because there were 28 respondents with more than five zero scores, as compared with only 10 who showed responsiveness using either of the two Interpretations above, for parity individuals are only listed as non-responsive in **Table 12** if they scored 8 or more zeros.

A poor response (rather than a non-response) could be considered as a 'bad' change – in the opposite direction to that considered healthy. If 'more than 2' bad scores was the criteria rather than '2 or more', then cases 8, 37, 48 and 54 could no longer be considered as poor responders. They could perhaps be classified as 'marginally poor responders'. Results for one of these (37) suggest she could be classified as both a 'good' and a 'marginally poor' responder.

In total, there were in total 317 MMCL zero scores, 120 '2' scores in the desired direction (Interpretation 1), or 114 using Interpretation 2, and 23 in the opposite direction (with only one person showing more than 2 items as 'bad' changes). This suggests that the MMCL, in this context, may be more unresponsive than it is responsive. However, poor responders would probably be unlikely among acupuncture aficionados who nearly all stated they benefited from acupuncture in the past, so the low number of 'poor responders' is not surprising.

Figure 2 shows unresponsiveness for the different MMCL items. Confused and Dispirited were the most unresponsive items, and Positive and Relaxed the least unresponsive.

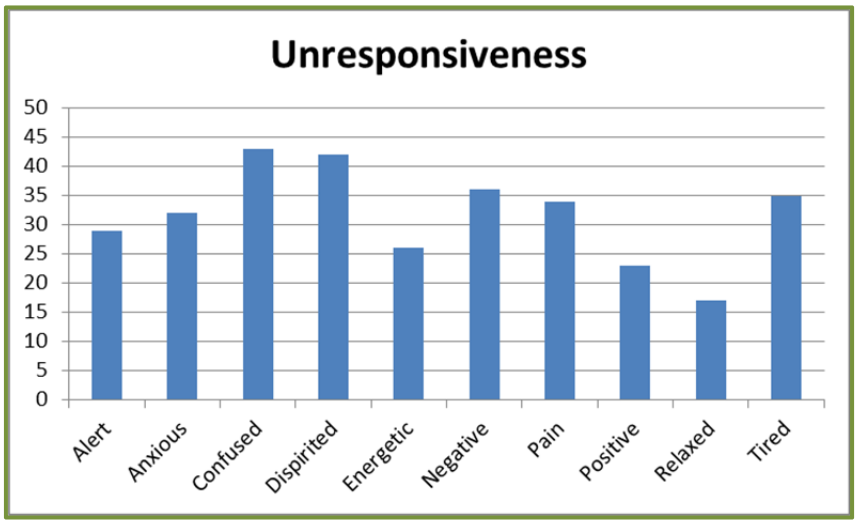

Fig 2. Unresponsiveness of the different MMCL items.

2.3. Shannon entropy (SE)

SE is a measure of the inherent uncertainty or randomness of information in a given string of data, defined as:

$$-\sum_{i=1}^{n} p(x_i)\log_2 p(x_i), \qquad \text{(E1)}$$

where X is the set of n categories under investigation and $p(x_i)$ is the probability of each x_i [Tastle & Wierman 2006]. It is sometimes also considered a measure of variability [e.g. Young & Wasserman 2001]. Surprisingly, given its origin in information theory, unlike RS (**Figure 1**), SE has rarely been used in the analysis of questionnaire responses, with only two relevant citations found in PubMed [Glynn et al. 2018; Handayani et al. 2017], although the longterm *dynamics* of questionnaire mood responses have been studied using another form of entropy, Approximate entropy [Pincus et al. 2008]. SE is used here as a prelude to further research in which entropy measures will be used to analyse both questionnaire responses and physiological measures.

In this study, SE was calculated using the open-access online calculator created by Lukasz Kozlowski (http://www.shannonentropy.netmark.pl/). A more rapid method for calculating SE in Excel is given in **Appendix A**.

For the different questionnaires used, SE and *normalised* SE (nSE, or SE divided by $\log_2 n$, where n is the number of items in the questionnaire[3]) in this study are shown in **Table** 13.

[3] This definition of nSE can be found on various websites [e.g. Anon. (2015). Tsallis and Rényi Normalized Entropy], but an academic reference for it could not be located.

Table 13. Shannon entropy (SE) and normalised SE for the questionnaires in this study.

Questionnaire	median SE (median nSE)	minimum SE	maximum SE
AHQ	1.371 (0.590)	0	2.521
TAS	1.561 (0.520)	0.544	2.446
BIS/BAS	1.613 (0.352)	0.175	2.017
BAQ	2.147 (0.515)	0.614	2.774
I-PANAS-SF	1.722 (0.518)	0	2.246
MMCL	1.359 (0.409)	0	2.446

There was no consistent or significant difference in median SE across all scales between younger and older attendees across all scales. For two scales (AHQ and TAS), SE was identical for younger and older attendees, and for three other scales (BAQ, I-PANAS-SF and MMCL) it was more for younger than for older attendees; only for BIS/BAS was it more for older than for younger respondents. There was a negative correlation between age and I-PANAS-SF (Spearman's *rho* = -0.210, significant at the 0.05 level), again suggesting that older respondents may demonstrate lower SE when completing some questionnaires. Given that complexity and variability of many disparate physiological measures tend to decrease with age [Manor & Lipsitz 2013], it would be of interest to investigate in a larger study whether questionnaire SE follows a similar pattern.

For all scales other than MMCL, SE was lower for men than for women. This is in keeping with the nonsignificant finding that MRS tends to be higher in men, and ERS in women.

Respondents whose SE was consistently in the lower or upper quartiles for those questionnaires ('consistently' low, for example, being interpreted as low in at least three questionnaires and never high) were as shown in **Table 14**.

Table 14. Respondents with consistently low or high SE across questionnaires.

Respondent (by number)	SE consistently low	SE consistently high
2[a]	0	1 [F, 40]
6	0	1 [F, 29]
7	0	1 [F, 32]
14	1 [F, 47]	0
31	1 [F, 35]	0
32	1 [F, 39]	0
34	1 [F, 49]	0
51	1 [M, 61]	0
76[a]	1 [F, 46]	0
79	1 [F, 28]	0
84[b]	1 [M, 38]	0
89[b]	1 [unknown]	0
90	1 [F, 39]	0
Totals [*N* F, M; median age]	10 [7 F, 2 M; 39]	3 [3 F, 0 M; 32]

a. ES low or high for four questionnaires; b. ES low for five questionnaires; the remainder scored low or high only for three.

Those cases in bold also appear in **Table 11** above (on MRS and ERS). For these particular individuals, low SE corresponds to MRS, and high SE to ERS, as would be expected.

2.3.1. Standard deviation (SD)

A more usual measure of variability or, more precisely, variation, is standard deviation (SD), the square root of 'variance' of a set of data. The standard deviation of questionnaire scores is a measure of 'response range', sometimes itself considered as a response style [Van Vaerenbergh & Thomas 2013]. Values of SD for the questionnaire scores are shown with those of median SE, median normalised MRS (MRS/*n*) and median normalised ERS (ERS/*n*) in **Figure 3**.

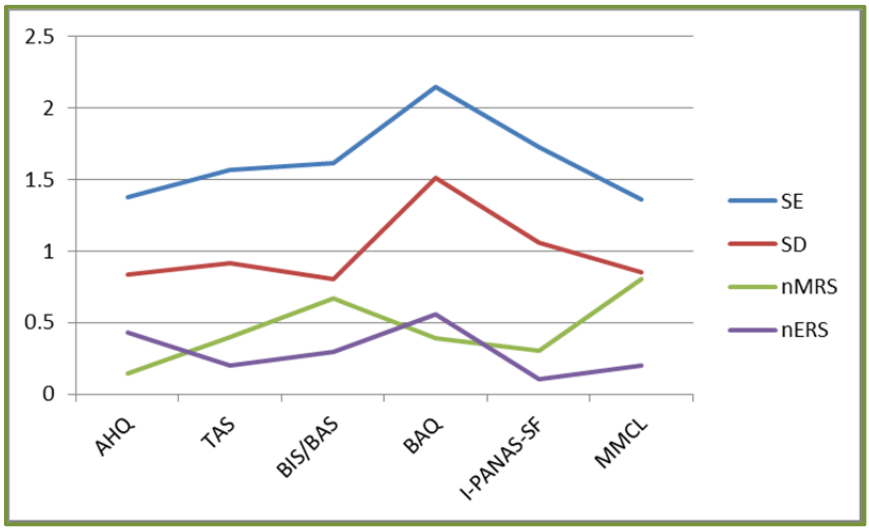

Fig 3. SE, SD, and normalised MRS and ERS for the primary questionnaires in this study.

The apparent similarity between SD and SE is discussed below (Section 13.1.1).

Hypotheses tested

A number of hypotheses were tested in this study. The results follow.

Hypothesis 1. The multiple mood change Likert scale (MMCL) shows acceptable validity and internal consistency

1.1. Internal consistency – Cronbach's alpha

Given that the different items in the MMCL were selected precisely because they were considered to measure different moods, it is not really meaningful to calculate Cronbach's *alpha* for the scale as a whole. Thus these *alpha* values will be low. This contrasts with findings for pre-to-post $BRUMS_{24}$ and NRS-M differences scores [Mayor & Steffert 2016] (**Table 15**).

Table 15. Cronbach's *alpha* for the $BRUMS_{24}$, NRS-M and MMCL scales.

***Alpha* values**	**$BRUMS_{24}$**	**NRS-M**	**MMCL**
Pre-treatment	0.694 (*N* = 167)	0.234 (*N* = 164)	
Post-treatment	0.634 (*N* = 165)	0.404 (*N* = 155)	
Pre-to-post treatment	0.813 (*N* = 145)	0.792 (*N* = 158)	0.137 (*N* = 60)

Low values of *alpha might* suggest better discriminant validity, but a more sensible approach would be to perform a cluster analysis initially, and then calculate Cronbach's *alpha* for the resulting clusters.

Using Ward's method and squared Euclidean distances in a hierarchical cluster analysis for a range of clusters between 2 and 5, the result is a very clear 2-cluster solution, separating the four moods that would be considered as therapeutically more 'desirable' from the six considered as more 'undesirable' (*N* = 60) (**Figure 4**).

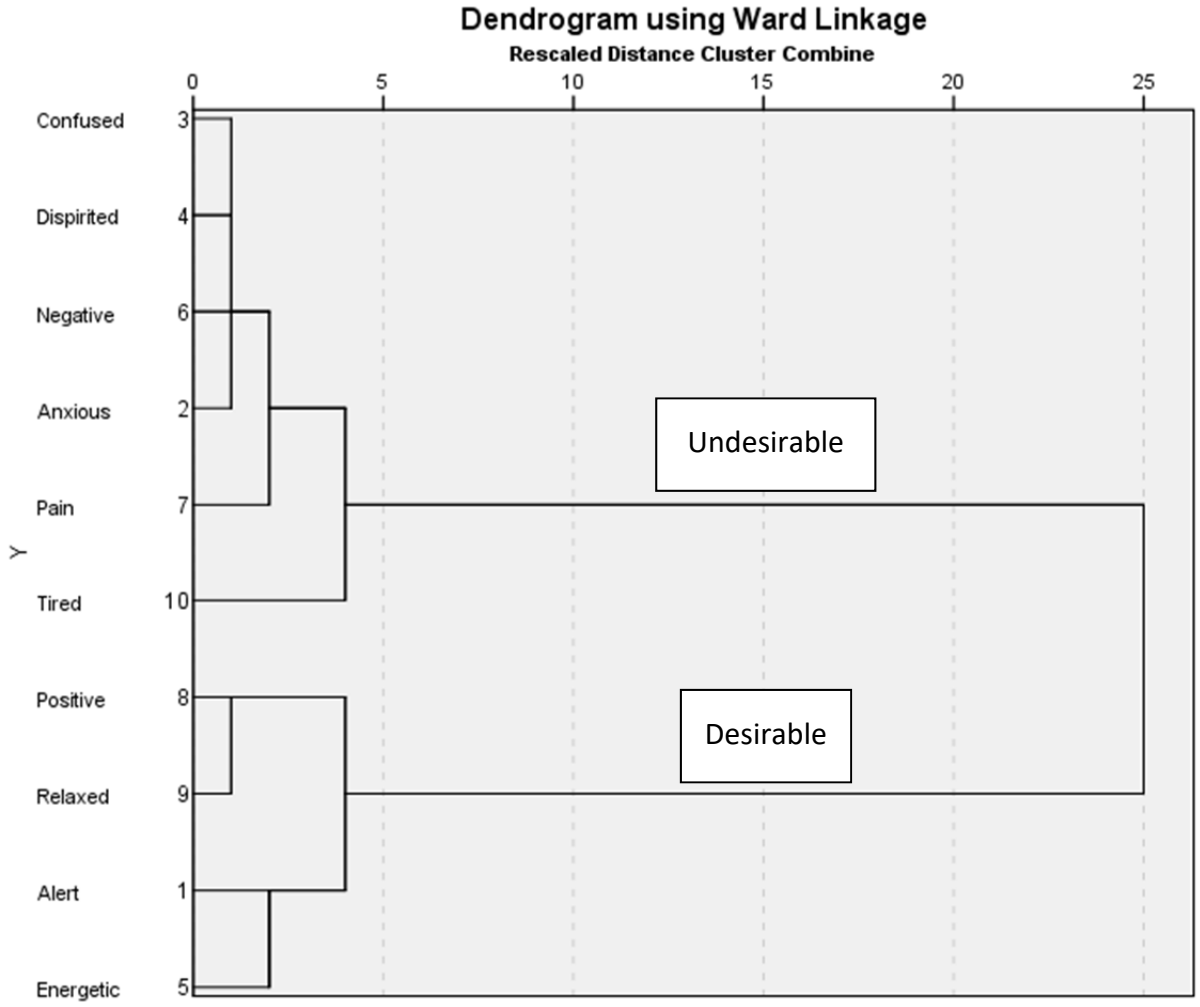

Fig 4. Results of hierarchical cluster analysis on MMCL items, showing how they fall naturally into 'desirable' and 'undesirable' clusters.

Even if 5 clusters are selected, four of the items remain clustered together in the 'undesirable' cluster, although the four in the 'desirable' cluster are then separated into two separate clusters – Positive/Relaxed and Alert/Energetic (**Table 16**).

Table 16. Results of cluster analysis, showing how MMCL items are distributed across different numbers of clusters.

Item	5 Clusters	4 Clusters	3 Clusters	2 Clusters
Alert	1	1	1	1
Anxious	2	2	2	2
Confused	2	2	2	2
Dispirited	2	2	2	2
Energetic	1	1	1	1
Negative	2	2	2	2
Pain	3	2	2	2
Positive	4	3	3	1
Relaxed	4	3	3	1
Tired	5	4	2	2

With Cronbach's *alpha* is calculated separately for the two 'desirable' and 'undesirable' mood clusters, values are more promising (**Table 17**).

Table 17. Cronbach's *alpha* for all MMCL items, and for 'desirable' and 'undesirable' mood items separately.

Scale	Cronbach's *alpha*	*N*
All mood items	0.137	60
'desirable' mood items	0.533	65
'undesirable' mood items	0.694	63

Thus there is more consistency among the 'undesirable' than the 'desirable' mood items.

However, the mean values for Alert and Energetic in the SPSS 'Item statistics' for *alpha* are very different (0.354 and 0.554, respectively) from those for Positive and Relaxed (both 0.923), suggesting again that a three-cluster solution may be preferable [Anon. n.d. Cronbach's alpha]. Recalculating *alpha* confirms this (**Table 18)**.

Table 18. Cronbach's *alpha* for a three-cluster solution.

Scale	Cronbach's *alpha*	*N*
All 'desirable' mood items	0.533	65
Positive/Relaxed mood items	0.684	68
Alert/Energetic mood items	0.558	67

1.2. Convergent and discriminant validity

Correlations between MMCL items were calculated using Spearman's *rho*, with the strongest correlations (those with absolute *rho* or $|rho| > 0.4$) and weakest correlations ($|rho| < 0.1$) shown in **Table 19**.

Table 19. Strongest and weakest positive and negative correlations between MMCL items. Correlations with *rho* > 0.5 are shown in **bold**.

Item	**Strongest correlations (|*rho*| > 0.4)**		**Weakest correlations (|*rho*| < 0.1)**	
	Positive	Negative	Positive	Negative
Alert	Energetic		Relaxed	Negative Pain
Anxious	**Confused** Dispirited Pain			
Confused	Anxious Dispirited			
Dispirited	Anxious Confused **Negative**			
Energetic	Alert Positive	Tired	Relaxed	
Negative	Dispirited			Alert
Pain	Anxious	Relaxed	Tired	Alert
Positive	Energetic Relaxed			Tired
Relaxed	**Positive**	Pain	Energetic	Tired
Tired		Energetic	Pain	Positive Relaxed

In particular, *rho* was > 0.5 only for the correlations between Anxious and Confused (0.544) , and between Dispirited and Negative (0.538) and Relaxed and Positive (0.538).

Convergent validity appears to be supported for the two (or three) clusters identified above, with Tiredness rather out on a limb on its own.

Divergent validity would be expected for several of the MMCL items, and indeed *rho* is low (and nonsignificant) between a number of these. As would be expected, there are strong, significant *negative* correlations (*rho* < 0.4) between Tired and Energetic, and between Pain and Relaxed.

In comparison, convergent validity (median |*rho*|) was similar for the NRS-M, but lower for $BRUMS_{24}$; divergent validity (|*rho*| < 0.1) was again similar for the NRS-M, but minimal for $BRUMS_{24}$.

Hypothesis 2. The MMCL is less sensitive to change in mood than pre-to-post differences in NRS-M or $BRUMS_{24}$ scores

Given the superior performance of NRS-M to $BRUMS_{24}$ when assessed using a variety of methods, although not all of them [Mayor & Steffert 2016], it was expected that the MMCL – which like $BRUMS_{24}$ is based on Likert scales – would again be less sensitive to changes in mood than the NRS-M.

However, the new mood change five-point Likert scales piloted appear – at least in the present context – to be quite sensitive to pre-to-post differences in mood.

Effect sizes (*r*) for pre-to-post NRS-M subscale differences in our previous classroom sample (*N* =158, using the Wilcoxon test for paired samples) are shown on the left in **Table 20** below [Mayor & Steffert 2016]. Effect sizes (Cohen's *d*) in the present, smaller, sample are shown on the right, based on a one-sample T-test for results differing from zero (no change). It can be seen that, where the moods measured are comparable, the effect sizes (ESs) are much greater using the Likert change scales – except for the Fatigue/Tired item.

Table 20. Sensitivity to change of the NRS-M (Wilcoxon test) and MMCL (one-sample T-test) for different samples, showing effect sizes.

NRS-M subscales (*N* = 134 of 158)	**p value**	**ES *r***	**MMCL mood changes (*N* = 68 of 90)**	**p value (2-tailed)**	**Cohen's *d* ($t/N^{1/2}$)**
Alert			Alert	0.006	0.35
Anxious	< 0.001	0.37	Anxious	< 0.001	0.64
Comfortable	ns	0.09	Comfortable		
Confused[a]	< 0.001	0.33	Confused	< 0.001	0.54
Fatigued	0.001	0.28	Tired	n.s.	0.05
Gloomy[a]	< 0.001	0.34	Dispirited	< 0.001	0.61
Lively	ns	0.10	Energetic	< 0.001	0.56
Pain			Pain	< 0.001	0.48
Relaxed	0.014	0.21	Relaxed	< 0.001	0.91
Overall positive mood	ns	0.14	Positive	< 0.001	1.10
			Negative	< 0.001	0.72
Median *r*		0.245	Median *d*		0.585

a. *N*=133

Figure 5 shows the average MMCL pre-to-post change for each mood. Note that changes are all in what might be considered a desirable direction, and all are significant (p < 0.01), except for the change in feeling Tired.

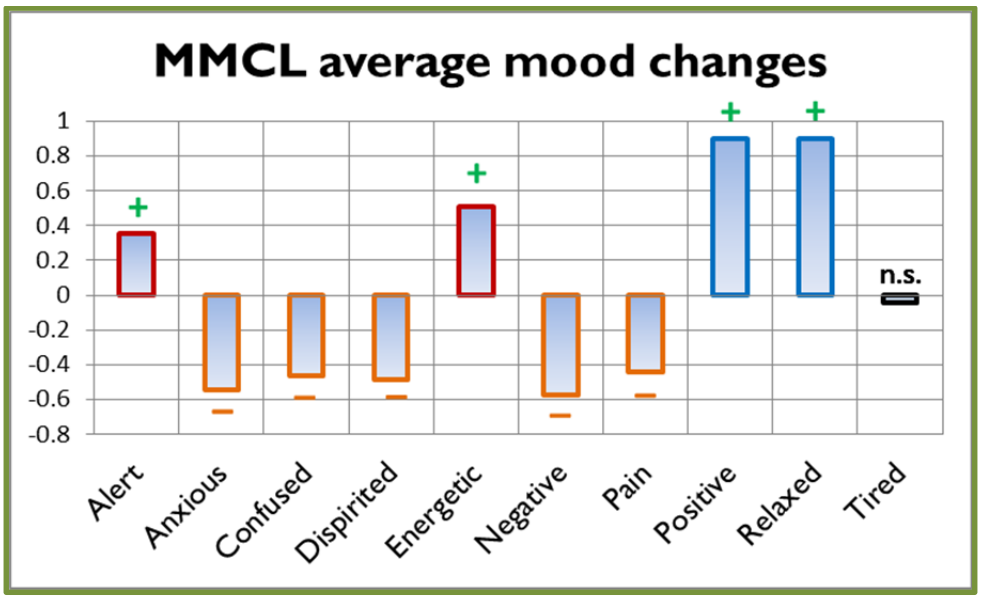

Fig 5. Average pre-to-post change for each mood.

When the percentage of *counts* of increases and decreases in NRS-M and MMCL mood change Likert scores in the two samples are compared, again p values are similar – or even somewhat smaller for the Likert change scores. In **Table 21** below, counts have been converted to percentages to assist comparison, and Relative Risk (RR) for changes in the desired direction is shown rather than ES, which is not appropriate for binary data [Anon. n.d. Effect size].

Table 21. Percentages of increases, decreases and no changes in mood using the NRS-M and MMCL on different samples, showing predominant directions of change and relative risk RR for binary data (instead of effect size as in the previous Table).

NRS-M subscale (*N* = 134 of 158)	% Inc	% Dec	% Zero	Binom p value	Rel Risk RR	Mood change (*N*=68 of 90)	% Inc	% Dec	% Zero	Binom p value	Rel Risk RR
Alert						Alert ↑	41.2	16.2	42.6	0.009	5.1
Anxious	18.7	55.2	26.1	< 0.001 ↓	5.9	Anxious ↓	7.4	45.6	47.1	< 0.001	12.4
Comfortable	43.3	37.3	19.4	n.s. ↑	2.3	Comfortable					
Confused[a]	25.6	49.6	24.8	0.00↓	3.9	Confused[b] ↓	4.5	31.3	64.2	< 0.001	14.0
Fatigued	28.4	61.2	10.4	< 0.001↓	4.3	Tired[c] ↓	21.7	27.5	50.7	n.s.	2.5
Gloomy[a]	19.5	52.6	27.8	< 0.001↓	2.9	Dispirited ↓	2.9	35.3	61.8	< 0.001	24.0
Lively	49.3	40.3	10.4	n.s. ↑	2.4	Energetic[c] ↑	52.2	10.1	37.7	< 0.001	10.3
Pain						Pain ↓	10.3	44.1	50.0	0.001	7.7
Relaxed	50.7	30.6	18.7	n.s. ↑	3.3	Relaxed[d] ↑	67.1	8.6	24.3	< 0.001	15.7
Overall pos mood	47.0	31.3	21.6	n.s. ↑	3.0	Positive ↑	64.7	1.5	33.8	< 0.001	88.0
						Negative ↓	2.9	44.1	52.9	< 0.001	30.0
Median RR					3.15	Median RR					13.2

a. *N* = 133; b. *N* = 67; c. *N* = 69; d. *N* = 70.

RR is clearly much greater for the mood change Likert scores than for the NRS-M differences, with more changes in what might be considered the desired direction for all the Likert scales (though not significantly so for feeling Tired), as shown in **Figure 6**.

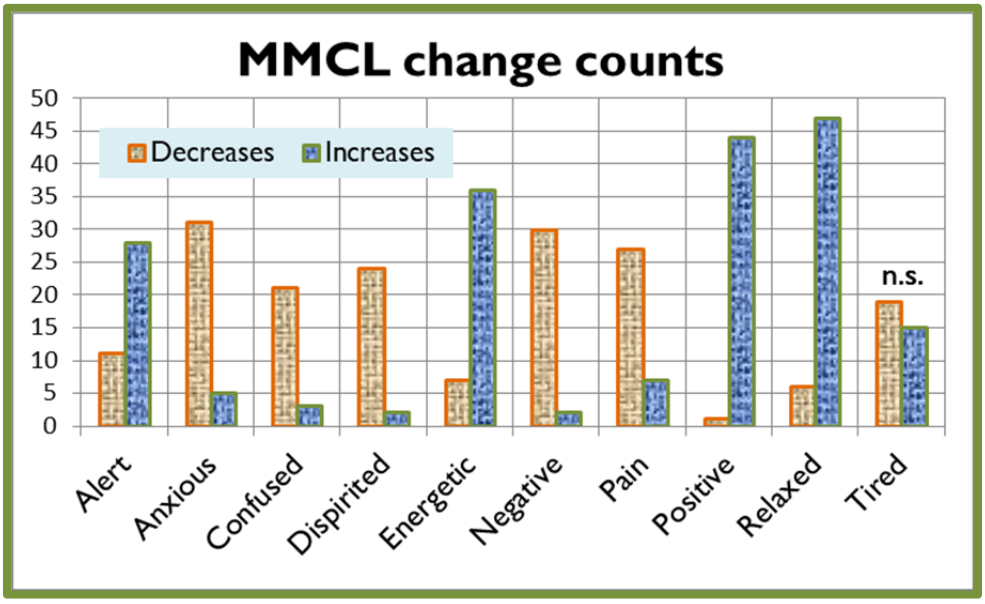

Fig 6. Numbers of increases and decreases in moods assessed using the MMCL.
Ratios of increases to decreases are all significant (p < 0.01) except for the change in feeling Tired.

Table 22 compares results for pre-to-post $BRUMS_{24}$ differences and MMCL mood change (both measures based on Likert scales), using one-sample T-tests for differences from zero, and with ES calculated as Cohen's *d*.

Table 22. Significance of pre-to-post $BRUMS_{24}$ differences and MMCL mood changes, using one-sample T-tests for differences from zero (ES shown as Cohen's *d*).

$BRUMS_{24}$ (*N* = 120 of 158)	**p value (2-tailed)**	**Cohen's *d***	**Mood change (*N* = 68 of 90)**	**p value (2-tailed)**	**Cohen's *d***
Alert			Alert ↑	0.006	0.35
Anger	< 0.001	0.38	Anger		
Tension	< 0.001	0.42	Anxious ↓	< 0.001	0.64
Confusion	< 0.001	0.44	Confused ↓	< 0.001	0.54
Fatigue	0.010	0.24	Tired ↓	n.s.	0.05
Depression	< 0.001	0.45	Dispirited ↓	< 0.001	0.61
Vigour	n.s.	0.04	Energetic ↑	< 0.001	0.56
Pain			Pain ↓	< 0.001	0.48
Relaxed			Relaxed ↑	< 0.001	0.91
Total mood disorder	< 0.001	0.42	Positive ↑	< 0.001	1.10
			Negative ↓	< 0.001	0.72
Median *d*		0.420	Median *d*		0.585

Apart from Tired/Fatigue, effect sizes are generally greater for the MMCL than for the $BRUMS_{24}$ scores. The former are predominantly medium or large, the latter are all small.

A comparison between percentages of *counts* of increases and decreases in $BRUMS_{24}$ and MMCL scores in the two samples is shown in **Table 23**.

Table 23. Percentages of increases, decreases and no changes in mood using $BRUMS_{24}$ and MMCL on different samples, showing predominant directions of change and relative risk RR for binary data.

$BRUMS_{24}$ (*N* = 120 of 158)	**% Inc**	**% Dec**	**% Zero**	**Binom p value**	**Rel Risk RR**	**Mood change (*N* = 68 of 90)**	**% Inc**	**% Dec**	**% Zero**	**Binom p value**	**Rel Risk RR**
Alert						Alert	41.2	16.2	42.6	0.009	5.1
Anger	10.8	37.5	51.7	< 0.001	6.9	Anger					
Tension	20.0	51.7	28.3	< 0.001	5.2	Anxious	7.4	45.6	47.1	< 0.001	12.4
Confusion	21.7	59.2	19.2	< 0.001	5.5	Confused	4.5	31.3	64.2	< 0.001	14.0
Fatigue	31.7	57.5	10.8	0.004	3.6	Tired	21.7	27.5	50.7	n.s.	2.5
Depression	7.5	40.8	51.7	< 0.001	10.9	Dispirited	2.9	35.3	61.8	< 0.001	24.0
Vigour	40.8	41.7	17.5	n.s.	2.0	Energetic	52.2	10.1	37.7	< 0.001	10.3
Pain						Pain	10.3	44.1	50.0	0.001	7.7
Relaxed						Relaxed	67.1	8.6	24.3	< 0.001	15.7
Total mood disorder	27.5	70.0	2.5	< 0.001	5.1	Positive	64.7	1.5	33.8	< 0.001	88.0
						Negative	2.9	44.1	52.9	< 0.001	30.0
Median RR					5.2						13.2

Again, apart from Tired/Fatigue, RRs are generally greater for the Mood change than for the $BRUMS_{24}$ scores.

However, for the MMCL, the percentage of zero (i.e., no change) scores is consistently far higher than for the NRS-M *or* $BRUMS_{24}$ differences, which is less encouraging (**Table 24**).

Table 24. Median percentage of no change scores for three different estimates of mood change.

NRS-M	$BRUMS_{24}$	MMCL
19.2%	20.5%	48.55%

Hypothesis 3. Some respondents may consistently report larger or smaller mood changes than the majority

As shown above in **Figures 5** and **6**, there were clearly some moods which changed more than others.

Many respondents recorded predominantly small changes in mood (0, or ±1), and some recorded large changes. The lowest possible sum of absolute changes over all 10 subscales was 0 (recorded by two respondents), the highest 20 (recorded by one respondent); thus the median score was 10.

Many more respondents scored below the median or the lower quartile (Q1) than above the median or upper quartile (Q3), as shown in **Table 25**. **Table 26** shows the total numbers of zero, ±1 and ±2 scores.

Paradoxically, although in several ways the MMCL appears more sensitive than pre-to-post differences in the NRS-M and $BRUM_{24}$ scores, the actual MMCL change scores are themselves mostly low (zero or ±1).

Table 25. Numbers of respondents and their sums of absolute change scores, by quartile.

Sums of absolute change scores							
Min (0)	<Q1 (5)	<median	Median (10)	>median	>Q3 (15)	Max (20)	**Total**
2	20	41	3	16	3	1	86

Table 26. Total numbers of zero, ±1 and ±2 scores.

***N* of 0, 1 and 2 absolute scores**		
Abs (2)	Abs (1)	Zeros
132	234	317

3.1. Skewness of response data

Skewness of questionnaire or scale response data may indicate that responses are biased in a particular direction. So when – as here – most respondents stated that they felt a lot more Relaxed after treatment, the results for this particular scale are 'negatively' skewed, whereas skewness for the change in feeling Tired after treatment is far smaller and 'positive' in direction (**Figure 7**). An absolute value of skewness divided by its standard error (SE) less than 1.96 is considered acceptable [Abu-Bader 2010]. In this case the skewness of Relaxed divided by its SE is -2.14 (so < -1.96), suggesting that Relaxed responses are significantly skewed, and not normally distributed, whereas for Tired it is +0.07, so more normally distributed. If all scale responses were unacceptably skewed in

the same direction, this could indicate that respondents were merely responding in the same way to all scales, regardless of what the scales purport to measure, but this is not the case here, although skewness is negative for eight items and positive for only two, with absolute values for the skewness of six out of the eight items being greater than the values of positive skewness.

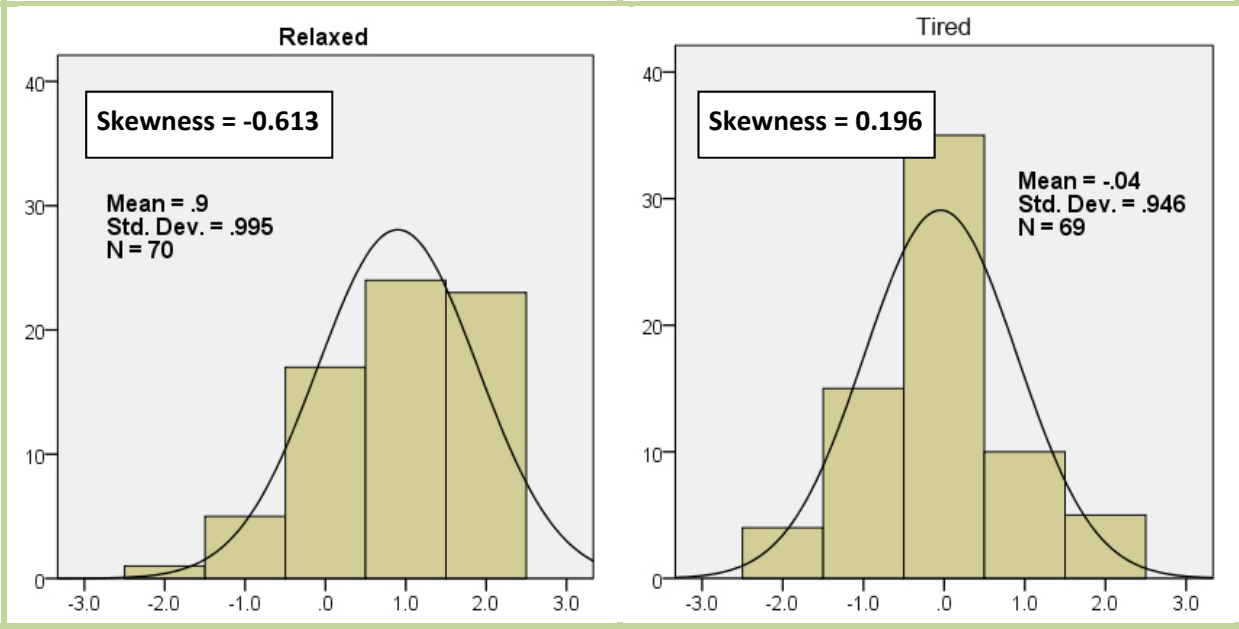

Fig 7. 'Relaxed' responses are negative skewed (Left), whereas those for 'Tired' (Right) are weakly positively skewed – more balanced between increases and decreases.

Median and maximum absolute skewness for the MMCL in this study are less than those for $BRUMS_{24}$ in our previous study. This is also true for median absolute skewness when MMCL is compared with NRS-M, although not for maximum absolute skewness. Overall, differences in median skewness again suggest that the MMCL may be more dependable than the other two measures, although maximum skewness/SE was higher in this sample than for NRS-M in our previous study (**Table 27**).

Table 27. Skewness for the MMCL compared to that for pre-to-post NRS-M and $BRUMS_{24}$ [Mayor and Steffert 2016]. Positive skewness indicated in red.

$BRUMS_{24}$ (*N* = 120 of 158)	**Skewness /SE**	**NRS-M subscale (*N* = 134 of 158)**	**Skewness/SE**	**MMCL (*N* = 68 of 90)**	**Skewness/SE**
Alert				Alert	-0.5
Anger	-5.4			Anger	
Tension	-2.4	Anxious	-2.5	Anxious	-1.0
		Comfortable	-1.6		
Confusion	-0.6	Confused	-1.2	Confused	-2.9
Fatigue	-0.2	Fatigued	1.4	Tired	0.7
Depression	-7.6	Gloomy	-2.6	Dispirited	-2.9
Vigour	1.6	Lively	-1.2	Energetic	-1.6
Pain				Pain	-0.6
		Relaxed	-2.1	Relaxed	-2.1
Total mood disorder	-1.8	Overall mood	-1.4	Positive	0.1
				Negative	-0.7
Median Skewness/SE	-1.8		-1.5		-1.31
Max Skewness/SE	7.6		2.6		2.9

3.2. Test discrimination $delta_G$

In our previous study [2016], we found that test discrimination, $delta_G$ (the ability of a scale to discriminate between individuals without reference to an external criterion [Hankins 2008]), was consistently greater for pre-to-post differences in NRS-M than for $BRUMS_{24}$, suggesting that NRS-M has greater sensitivity.

The equation for $delta_G$ is:

$$\delta = \frac{(1+k(m-1))(n^2-\sum_i f_i^2)}{n^2 k(m-1)} \qquad \text{(E2)}$$

Where *k* is the number of items (here, 1 if each scale is considered separately, or 10 if taken together), *n* is sample size (here, 67 to 70, depending on item, or their sum for the complete MMCL), *f* is the frequency of each score, and *m* is the number of item responses (here, 5).

$Delta_G$ values for pre-to-post $BRUMS_{24}$ and NRS-M, and for MMCL, are shown in **Table 28**.

Table 28. *Delta*$_G$ values for pre-to-post BRUMS$_{24}$ and NRS-M, and for MMCL mood changes

BRUMS$_{24}$ (*N* = 120 of 158)	***delta*$_G$**	**NRS-M subscale (*N* = 134 of 158)**	***delta*$_G$**	**MMCL (*N* =68 of 90)**	***delta*$_G$**
Alert				Alert	0.9099
Anger	0.6946			Anger	
Tension	0.8173	Anxious	0.8614	Anxious	0.8256
		Comfortable	0.8431		
Confusion	0.8461	Confused	0.8774	Confused	0.6677
Fatigue	0.8973	Fatigued	0.9167	Tired	0.8362
Depression	0.7007	Gloomy	0.8799	Dispirited	0.6937
Vigour	0.8610	Lively	0.9066	Energetic	0.8549
Pain				Pain	0.8253
		Relaxed	0.8311	Relaxed	**1.0110**
Total mood disorder	0.8307	Overall mood	0.7817	Positive	**0.9143**
				Negative	0.7683
		TOTAL	0.9037	**TOTAL**	0.7929

Median values of *delta*$_G$ for the three measures are, respectively, 0.8307, 0.8694 and 0.8309. Whereas for comparable items in NRS-M and BRUMS$_{24}$ *delta*$_G$ is generally higher for NRS-M, this is not the case for MMCL, for which only two items show higher *delta*$_G$ than the corresponding items in NRS-M, and only one item for the corresponding item in BRUMS$_{24}$.

Thus, MMCL discriminates best between individuals on the 'desirable' mood items (range 0.85 to 1.01, median 0.91), and somewhat less well on the 'undesirable' mood items (range 0.67 to 0.84, median 0.80).

Hypothesis 4. Mood changes found, and their magnitude and direction, may be affected by personality type or attitude

Table 29 shows correlations between MMCL results and the various other primary measures.

Table 29. Associations between MMCL results and the other primary measures.

Questionnaire (Spearman's *rho*)	TAS (pos or neg)	BIS/BAS	BAQ	I-PANAS-SF (pos or neg)	JTT (Mann-Whitney)
Alert	n.s.	BAS-F: -0.267 p=0.028 BIS: 0.294 p=0.015	n.s.	n.s.	n.s.
Anxious	n.s.	n.s.	-0.245 p=0.044	P-N: -0.303 p=0.013	n.s.
Confused	n.s.	n.s.	n.s.	n.s.	n.s.
Dispirited	T-P: 0.259 p=0.033	n.s.	-0.315 p=0.009[a]	n.s.	n.s.
Energetic	n.s.	n.s.	n.s.	n.s.	n.s.
Negative	n.s.	n.s.	n.s.	n.s.	n.s.
Pain	n.s.	n.s.	n.s.	n.s.	n.s.
Positive	n.s.	n.s.	n.s.	n.s.	n.s.
Relaxed	n.s.	n.s.	0.245 p=0.041	P-P: 0.352 p=0.003[a]	n.s.
Tired	n.s.	n.s.	n.s.	n.s.	n.s.

a. See too below, **Table 36**.

TAS. There appears to be some association between a positive attitude to technology and feeling more Dispirited after EA. This is not easy to interpret or particularly convincing.

BIS/BAS. There appears to be a negative association between Fun seeking and feeling less Alert after EA, and a positive association between behavioural inhibition and feeling more Alert after EA. These findings are not immediately understandable, but may merit future investigation – particularly the latter.

BAQ. Being less bodily aware may result in feeling more Anxious after EA, and being more bodily aware may result in feeling more Relaxed after EA. These findings are intriguing, even though the correlations are not strong, and again may merit future investigation. The association between being more bodily aware and feeling less Dispirited after EA is stronger, and also suggests further avenues to explore.

I-PANAS-SF. A generally less negative mood is associated with feeling more Anxious following EA, and a generally more positive mood with feeling more Relaxed after EA. The first of these is difficult to understand, but the second makes some sense. Again, these could be investigated further.

JTT. Rather as in our previous study on NRS-M [Steffert & Mayor 2017], here median values of all MMCL items were not significantly different for any JTT type (using the Mann-Whitney test), although following treatment Introverts tended to feel more Alert and Energetic, and less Anxious, than Extraverts. This is a finding that may merit further investigation.

Hypothesis 5. Benefit from prior acupuncture as assessed using the AHQ may be reflected in mood changes found during the teaching session

Only small, if occasionally significant, correlations were found between changes in feeling following treatment during the teaching session and past helpfulness of acupuncture, as follows:

5.1. Anxious and past helpfulness for Functional/unexplained disorders (*rho* = -0.261, p = 0.049, *N* = 57)

5.2. Relaxed and past helpfulness for Physical/musculoskeletal disorders (*rho* = 0.248, p = 0.048, *N* = 64)

5.3. Confused and generally helpful in the past (*rho* = 0.246, p = 0.046, *N* = 66).

However, the effect sizes for these are only small (*rho* ~ ES), so may be discounted as likely to be the result of chance.

Hypothesis 6. Benefit from prior acupuncture may be reflected in EA stimulation parameters described as most or least pleasant and/or intense

Hypothesis 7. Which EA stimulation parameters are described as most or least pleasant and/or intense may be affected by personality type or attitude

Hypothesis 8. MMCL scores may show an association with which stimulation parameters are found to be most or least pleasant and/or intense

These hypotheses are considered together.

Using the Mann-Whitney test to check if there were any differences in prior acupuncture experience, personality type or attitude, or MMCL scores, for the two stimulation parameters LF and HF, none were found, other than for TAS (positive attitude to technology) (**Table 30**).

Those who stated that they felt LF stimulation to be least intense had a more positive attitude to technology (median 3.5) than those who felt HF stimulation to be least intense (median 2.75). Given the number of comparisons undertaken, the small-to-medium effect size *r*, and the difficulty in interpreting this result with any confidence, it can probably be discounted as due to chance.

Table 30. Significance of differences in questionnaire responses between those who considered LF or HF stimulation parameters most or least pleasant/intense.

Questionnaire	Most intense	Least intense	Most pleasant	Least pleasant
AHQ	n.s.	n.s.	n.s.	n.s.
TAS	n.s.	p=0.035[a], *U* = 49.5; *r* = 0.40	n.s.	n.s.
BIS/BAS	n.s.	n.s.	n.s.	n.s.
BAQ	n.s.	n.s.	n.s.	n.s.
I-PANAS-SF	n.s.	n.s.	n.s.	n.s.
MMCL	n.s.	n.s.	n.s.	n.s.

a. Asymptotic significance.

For the JTT, counts for each combination of type and whether LF or HF was entered did not show any significant differences between types (using the Binomial test) (**Table 31**).

Table31. Counts of JTT types reporting LF or HF as most/least intense or pleasant.

Questionnaire	*N*	Most intense	Least intense	Most pleasant	Least pleasant	Totals
Introversion (I)	36	5 LF; 8 HF	9 LF; 4 HF	13 LF; 4 HF	4 LF; 11 HF	31; 27
Extraversion (E)	34	8 LF; 11 HF	6 LF; 4 HF	4 LF; 5 HF	7 LF; 7 HF	25; 27
Intuition (N)	45	8 LF; 12 HF	11 LF; 6 HF	9 LF; 6 HF	6 LF; 10 HF	34; 34
Sensing (S)	25	5 LF; 7 HF	4 LF; 2 HF	8 LF; 3 HF	5 LF; 8 HF	22; 20
Feeling (F)	57	11 LF; 13 HF	13 LF; 7 HF	14 LF; 8 HF	8 LF; 13 HF	46; 41
Thinking (T)	13	2 LF; 6 HF	2 LF; 1 HF	3 LF; 1 HF	3 LF; 5 HF	10; 13
Perceiving (P)	39	6 LF; 6 HF	6 LF; 6 HF	8 LF; 5 HF	5 LF; 9 HF	25; 26
Judging (J)	31	7 LF; 13 HF	9 LF; 2 HF	9 LF; 4 HF	6 LF; 9 HF	31; 28
Totals ÷ 4	70	13 LF; 19 HF	15 LF; 8 HF	17 LF; 9 HF	11 LF; 18 HF	54; 56

Note: Because of missing data, these figures do not tally with those in **Table 9**.

Hypothesis 9. Whether acupuncture is reported as beneficial in the past may depend on personality type or attitude

There were significant positive correlations between AHQ items only for PANAS-Pos, as shown below. No other correlations were significant (**Table 32**).

Table 32. Correlations between AHQ items and questionnaires on personality type or attitude, using Spearman's *rho*.

Questionnaire	Helpful	Physical	Functional	Emotional	Lifestyle
TAS	n.s.	n.s.	n.s.	n.s.	n.s.
BIS/BAS	n.s.	n.s.	n.s.	n.s.	n.s.
BAQ	n.s.	n.s.	n.s.	n.s.	n.s.
I-PANAS-SF	n.s.	P-Pos: 0.356**	P-Pos: 0.252*	n.s.	P-Pos: 0.359**

There were differences in the median scores for some of the AHQ items between several of the JTT dyads, although only one of these approached significance (**Table 33**).

Table 33. Comparison of median AHQ item scores for the four JTT type dyads .

Questionnaire	Helpful	Physical	Functional	Emotional	Lifestyle
JTT	I = E N = S F = T P = J	I = E N > S F > T P = J	E > I (ns) N < S F > T P < J	I = E N > S (p=0.059) F > T P=J	I = E N = S F > T P = J

Thus those with a more positive general mood are likely to report benefits from past treatment, but otherwise personality, attitude or Jungian type did not appear to affect reporting of past benefits from acupuncture. In particular, there was no difference between Feeling and Thinking types for overall helpfulness of past acupuncture – in contrast to one of the results from our previous study, in which Thinking types stated that they considered acupuncture as more helpful (Mann-Whitney U test p = 0.003, *r* = 0.38) [Steffert & Mayor 2017].

Hypothesis 10. There may be differences in personality type and/or attitude among training institutions

In our previous report [Steffert & Mayor 2017 poster], some differences were found. A Kruskal-Wallis test was carried out to test the hypothesis on the present data. Results are summarised in **Table 34**.

Table 34. Results of a Kruskal-Wallis test comparing differences in questionnaire responses among training institutions, and showing for which institutions median scores were high or low. MRS = Midpoint response style; SE = Shannon entropy.

Questionnaire	Kruskal-Wallis	AA	CICM	LSBU	NCA
AHQ	AHQ-MRS: p=0026	*lowest* medians on 4 items			*highest* medians on 4 items
TAS		TAS-Neg lowest		TAS-Pos highest	TAS-Pos lowest; TAS-Neg highest
BIS/BAS	BAS-Drive: p=0.001 BB-MRS: p=0.040		BIS highest	BAS-Drive highest BAS-Fun lowest BAS-Reward highest BIS lowest	BAS-Drive lowest
BAQ	BAQ-MRS: p=0.034	lowest			highest
I-PANAS-SF		P-Pos high P-Neg low	P-Pos high	P-Pos low	P-Pos low
JTT		See **Section 1.6** above for JTT differences			
MMCL	Negative: p=0.048 Pain: p=0.048 MMCL-SE: p=0.012	Neg: lowest		Pos: lowest	Alert: highest Anx: lowest Disp: lowest Energ: lowest Neg: lowest Pain: lowest

Analysis by paired Institutions, using the Mann-Whitney test, showed some significant differences (**Table 35**).

Table 35. Significant differences in questionnaire scores between paired training institutions.

Questionnaire	Mann-Whitney	AA	CICM	LSBU	NCA
AHQ	Functional: p = 0.010, *U* = 77.0, *Z* = -2.580, *r* = 0.42	*			*
TAS	TAS-Neg: p = 0.013, *U* = 123.5 *Z* = -2.490, *r* = 0.38	*			*
BIS/BAS	BAS-Drive: p < 0.001, *U* = 18.5 *Z* = -3.598, *r* = 0.60 BAS-Drive: p = 0.002, *U* = 45.5 *Z* = -3.084, *r* = 0.46		 **	*** **	***
BAQ	n.s.				
I-PANAS-SF	n.s.				
MMCL	Pain: p = 0.012 *U* = 183.5, *Z* = -2.506, *r* = 0.36 Tired: p = 0.018; *U* = 200.5 *Z* = -2.366, *r* = 0.33 Negative: p = 0.022, *U* = 12.5 *Z* = -2.299, *r* = 0.53	 *	* * 	 *	* *

* Significant at the 0.05 level; ** significant at the 0.01 level; *** significant at the 0.001 level.

Median values of BAS-D were lowest for the NCA, and highest for LSBU.

Hypothesis 11. Missing data may depend on personality type, attitude, or training institution

See Section on Missing Data above (pp 7-8) for results and brief discussion.

Hypothesis 12. Other significant interactions between primary measures, between primary and secondary measures, and between secondary measures

12.1. Other interactions between primary measures

There were only two values of Spearman's *rho* significant at the 0.01 level (i.e. with *rho* greater than around 0.310) for correlations between the primary measures and MMCL mood changes (**Table 36**, and compare **Table 29**, above).

Table 36. Significant associations between MMCL mood changes and personality measures.

Personality scale	MMCL item	*Rho* (p)	Interpretation
BAQ	Dispirited	-0.315 (p=0.009)	Feeling less Dispirited after treatment may be associated with greater bodily awareness
I-PANAS-SF Positive	Relaxed	+0.352 (p=0.003)	Feeling generally more Positive in life may be associated with increased Relaxation after treatment

There were several correlations significant at the 0.01 level among the primary measures themselves – however, these were mostly as would be expected. The most interesting were those between BAS-Drive and PANAS positive affect, and BIS and PANAS negative affect (**Table 37**).

Table 37. Significant associations among primary measures, using Spearman's *rho*.

rho	TAS +	TAS -	BAS-D	BAS-F	BAS-R	BIS	BAQ	PANAS+	PANAS-
TAS +		-0.353 p=0.001							
TAS -									
BAS-D				0.387 P<0.001	0.484 P<0.001			**0.405** **P<0.001**	
BAS-F					0.387 P<0.001				
BAS-R									
BIS									**0.455** **P<0.001**
BAQ									
PANAS+									
PANAS-									

Mann-Whitney tests were performed for the various scales, together with their response type and SE, with the four JTT type dyads as grouping variables (24 comparisons for each dyad). No results were significant for the Sensing/Intuition or Perceiving/Judging dyad; there were three others that were, but only at the 0.05 level (**Table 38**).

Table 38. Significant differences in personality measures between JTT dyads.
Note the small effect sizes.

Scale	JTT: I vs E	JTT: F vs T
TAS neg	*U* = 446.500 *Z* = -1.965 p = 0.049; *r* = 0.23	
BAS-Drive	*U* = 444.500 *Z* = -2.182 p = 0.029; *r* = 0.26	
BAS-Fun		*U* = 250.000 *Z* = -2.148 p = 0.032; *r* = 0.25

For both TAS (neg) and BAS-Drive, median scores were higher for Extraverts than Introverts; for BAS-Fun, median scores were higher for Feeling rather than Thinking types.

These results all have a small effect size (*r*), but may still be worth investigating further. For example, it makes some sense that BAS-Drive may be higher in Extraverts than Introverts, and the median scores for all BAS subscales were indeed greater for Extraverts, if not significantly so.

An average BAS score, BAS-av, was created from the three separate BAS scores, although the three BAS subscales do not measure the same construct. Nonetheless, Extraverts scored higher than Introverts on BAS-av, the difference only just missing significance (**Table 39**).

Table 39. The relationship between BIS/BAS and JTT Introversion/Extraversion in this sample (Mann-Whitney test).

Median scores	Introvert	Extravert	Significance
BAS-av	12.67	13.33	p=0.0505 (*r* = 0.23)
BIS	19.00	18.00	n.s. (*r* = 0.07)

Thus the Introvert/Extravert and BIS/BAS-av constructs are not really comparable, and in any future research it would not therefore be possible simply to replace the Jungian Introvert/Extravert dyad with BIS/BAS.

A hierarchical cluster analysis was performed as above on 'high' vs 'low' scores (defined relative to median score) for the various personality/attitude scales considered together. Interestingly, for a 2-cluster solution, the scales neatly divided into what might be considered 'undesirable' and 'desirable' clusters (**Table 40**).

Table 40. Results of a hierarchical cluster analysis on high and low personality and attitude scale scores.

'Undesirable' cluster	'Desirable' cluster
BIS I-PANAS-SF negative TAS negative	BAQ BAS Drive BAS Reward BAS Fun I-PANAS-SF positive TAS positive

For a 4-cluster solution, the two clusters subdivided, with those scales underlined in the Table above forming single-scale clusters on their own.

12.2. Other interactions between primary and secondary measures

12.2.1. Correlations using Spearman's *rho* are shown in **Table 41**, with primary measures in the first column and the secondary measures after the slash (/) in the other columns.

Table 41. Significant correlations between primary and secondary measures, using Spearman's *rho*.

Questionnaire	MRS	ERS	SE
AHQ	**gnrl/PANAS -0.332****	***gnrl/ALL -0.315****	phys/AHQ -0.344**
	gnrl/ALL -0.292*	phys/TAS 0.254*	phys/PANAS 0.258*
	phys/TAS 0.254*	phys/BAQ 0.229*	func/AHQ -0.449***
	phys/PANAS -0.281*	**phys/PANAS 0.305****	func/MMCL 0.232*
	phys/ALL -0.391**	phys/MMCL 0.246*	emot/TAS 0.223*
	func/BB -0.234*	***phys/ALL -0.438*****	lifestyle/AHQ -0.349**
	func/ALL -0.409**	func/MMCL 0.264*	
	lifestyle/ALL -0.281*	***func/ALL 0.375****	
		emot/TAS 0.246*	
TAS	TAS -0.320**	TAS 0.274**	
BIS/BAS	BAS-D BB -0.446***	BAS-D BB 0.374***	BAS-D BB -0.213*
	BAS-D PANAS -0.251*	**BAS-D PANAS 0.326****	BAS-D MMCL -0.233*
	BAS-D ALL -0.351*	BAS-D ALL 0.264*	
	BAS-F BB -0.426***	BAS-F BB 0.403***	BAS-F BB -0.302**
	BAS-F ALL -0.285*	***BAS-F ALL -0.330****	
	BAS-R BB -0.601***	BAS-R BB 0.558***	
	BAS-R ALL -0.378**	BAS-R PANAS 0.243*	
	BIS BB -0.247*	***BAS-R ALL 0.358****	
	BIS BAQ 0.219*	BIS BB 0.302**	
	BIS PANAS 0.310**	BIS PANAS -0.264*	
BAQ	BAQ -0.300**	BAQ 0.325**	BAQ -0.332**
	PANAS -0.322**	**MMCL 0.322****	**MMCL 0.309****
	ALL -0.218*	***ALL 0.288****	
I-PANAS-SF	Pos PANAS -0.628***	Pos BB 0.213*	Pos AHQ -0.213*
	Pos ALL -0.400**	Pos PANAS 0.569***	Pos PANAS 0.293**
	Neg PANAS 0.612***	***Pos ALL 0.340****	
		Neg PANAS -0.444***	Neg PANAS -0.283**
MMCL	Alert PANAS 0.283*	Alert PANAS -0.250*	Anxious AHQ 0.256*
	Alert MMCL -0.248*		Anxious BAQ 0.285*
	Anxious MMCL 0.609***	Anxious MMCL -0.633***	Anxious MMCL -0.541***
	Confus MMCL 0.549***	Confus MMCL -0.475***	Confus MMCL -0.318**
	Dispir MMCL 0.763***	Dispir MMCL -0.516***	Dispir MMCL -0.494***
	Energ BAQ 0.238*		
	Energ PANAS 0.275*		Energ PANAS -0.268*
	Energ MMCL -0.387**	Energ MMCL 0.314**	
	Neg MMCL 0.618***	Neg MMCL- 0.530***	Neg MMCL -0.429***
	Pain MMCL 0.416***	Pain MMCL -0.491***	Pain MMCL -0.358**
			Pos BB -0.263*
	Pos MMCL -0.404**	Pos MMCL 0.663***	Pos MMCL 0.383**
			Relax TAS 0.241*
		Relax PANAS 0.246*	Relax PANAS 0.262*
	Relax MMCL -0.404**	Relax MMCL 0.692***	Relax MMCL 0.440***
		Relax ALL 0.248*	
	Tired MMCL -0.280*		
			Tired BB -0.247*

* significant at the 0.05 level; ** significant at the 0.01 level; *** significant at the 0.001 level.

Discounting the significant correlations for scales with themselves, and those with 'ALL', there remain a few correlations significant at the 0.001 level or better: AHQ subscales with PANAS (MRS or ERS), BIS or BAS-Drive with PANAS (MRS or ERS), BAQ with PANAS (MRS) and MMCL (ERS and SE).

All these make intuitive sense, and several suggest that it may be worth continuing investigations of the I-PANAS-SF in further research.

12.2.2. JTT and the secondary measures

Mann-Whitney tests were performed for SE and ERS for the various scales, with the four JTT type dyads as grouping variables (24 comparisons for each dyad). Two results were significant for the Perceiving/Judging dyad, but only at the 0.05 level; there were no others that were (**Table 42**).

Table 42. Significant differences in personality and attitude scale SE and ERS values for the JTT dyads.

Scale	JTT: P vs J
TAS-SE	*U* = 422.500 *Z* = -2.344 p = 0.019; *r* = 0.28
TAS-ERS	*U* = 434.000 *Z* = -2.327 p = 0.029; *r* = 0.28

For TAS SE and ERS, median scores were higher for Perceiving than for Judging, and it is intriguing that there may apparently be a greater variability in TAS responses for Perceivers than for Judgers.

Nonsignificant difference in median RS and SE scores between Perceivers and Judgers are summarised in **Table 43**.

Table 43. Nonsignificant difference in median RS and SE scores between Perceivers and Judgers, indicating agreement (or lack of it) in sign of J-P difference in MRS, ERS and SE scores.

Questionnaire	MRS	ESR	SE	
AHQ	J > P	J < P	J < P	agreement
TAS	J = P	J < P	J < P	[agreement]
BIS/BAS	J < P	J > P	J < P	disagreement
BAQ	J > P	J < P	J < P	agreement
I-PANAS-SF	J = P	J = P	J < P	[agreement]
MMCL	J > P	J = P	J > P	[disagreement]
ALL scales	J > P	J < P	na	agreement

More scales demonstrate agreement than disagreement between the secondary measures MRS, ERS and SE. Thus it may be useful to investigate this further, but only with a large sample of respondents.

12.2.3. Responsiveness and primary measures.

Numbers of 'good', 'poor' and 'non-' responders in **Table 12** above are small, but a Mann-Whitney test was carried out nonetheless for all measures to see if there was a significant difference between 'good' and 'non' responders. Given that responsiveness was defined by MMCL changes, significant differences were found for most MMCL items (except for Energetic and Tired), and for MMCL SE and

RS, but not for any other measure except for helpfulness of prior acupuncture for functional conditions: those who found it helpful were more likely to be non-responders. This finding does not appear likely, and was significant only at the 0.05 level (U = 7.5, Z = -2.079, p = 0.042, N = 14, *r* = 0.56; *r* is inflated because of the small sample size).

Median scores were compared for nonsignificant differences, but none of those between good and non-responders made particular sense. Good responders did, however, show higher median scores than the few poor responders for general helpfulness of prior acupuncture and its helpfulness for functional and emotional conditions, as well as lifestyle advice given, for all BIS/BAS subscales except BAS-Fun, for BAQ and for PANAS Positive (with poor responders scoring higher than good responders on PANAS Negative). These mostly make intuitive sense, but verification would require a much larger sample.

Hypothesis 13. Some respondents may show similar patterns of responsiveness, response style or Shannon entropy over a number of the mood scales and/or personality/attitude measures

Hypothesis 14. There may be significant interactions among the secondary measures themselves

These hypotheses are considered together.

13/14.1. Response style

Hypothesis 13.1.1. Individuals may complete all questionnaires in a similar way, as assessed by response style

This does not appear to be the case in general, although there is a certain amount of agreement between response style for *some* of the questionnaires (**Table 44**).

Table 44. Correlations among response styles for the different questionnaires, using Spearman's *rho.*

rho	AHQ	TAS	BB	BAQ	PANAS	MMCL
AHQ-M	-***^{E}		**^{M}		**^{M}	
AHQ-E			**^{E}		**^{E}	
TAS-M		-**^{E}	*^{M}			
TAS-E			*^{E}	*^{E}	*^{E}	
BB-M			-***^{E}	*^{M}	-**^{E}	
BB-E				*^{E}	**^{E}	
BAQ-M				-***^{E}	-**^{E}	
BAQ-E					**^{E}	
PANAS-M					-***^{E}	
PANAS-E						
MMCL-M						-***^{E}
MMCL-E						

*** = correlation significant at the 0.001 level; ** = significant at the 0.01 level;
* = significant at the 0.05 level; + = positive correlation; - = negative correlation;
E = ERS for questionnaire in that column; M = MRS for the questionnaire.

Hierarchical cluster analysis of response style, using the same method as before (and omitting 'ALL') indicates a 2-cluster solution, with BIS/BAS and BAQ in one cluster, and the other measures in the other cluster (**Figure 8**).

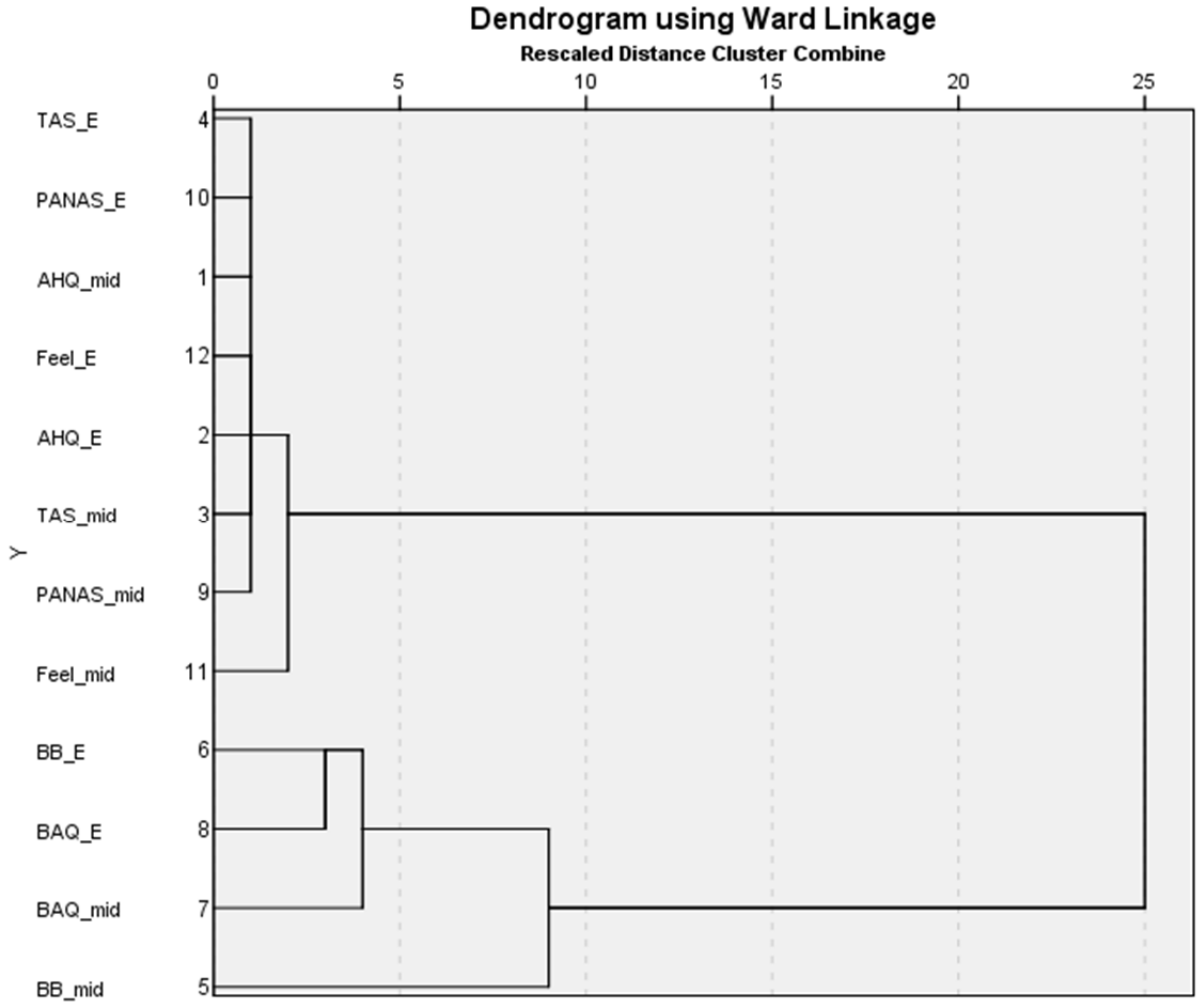

Fig 8. Results of hierarchical cluster analysis on Response style, showing how the questionnaires fall into two clusters, with BIS/BAS and BAQ in one cluster, and the other measures in the other cluster.

To further explore the relationship between MMCL changes and personality or attitude, a cluster analyses of *cases* (rather than variables) was then attempted, for the response styles of all the primary questionnaires together with each of the MMCL items in turn. For most of these, a 2-cluster solution was possible (if not necessarily optimal for Alert or Positive, for example). Respondents tended to be in the same 58 clusters for both Positive and Negative, or for both Anxious and Relaxed, with only a few individuals in the different clusters (9 and 10, respectively). More detailed exploratory analysis along these lines was not conducted, partly because of the amount of missing MMCL data, and partly because potential benefits would probably not justify the work involved.

As would be expected, there are strong negative correlations between Midpoint and Extreme response styles for the *same* questionnaires (varying between *rho* = -0.994 for BAQ and -0.968 f0r BIS/BAS ($p<0.001$) to *rho* = -0.297 for TAS ($p = 0.004$), with -0.752 for all questionnaires considered together ($p<0.001$) (**Figure 9**).

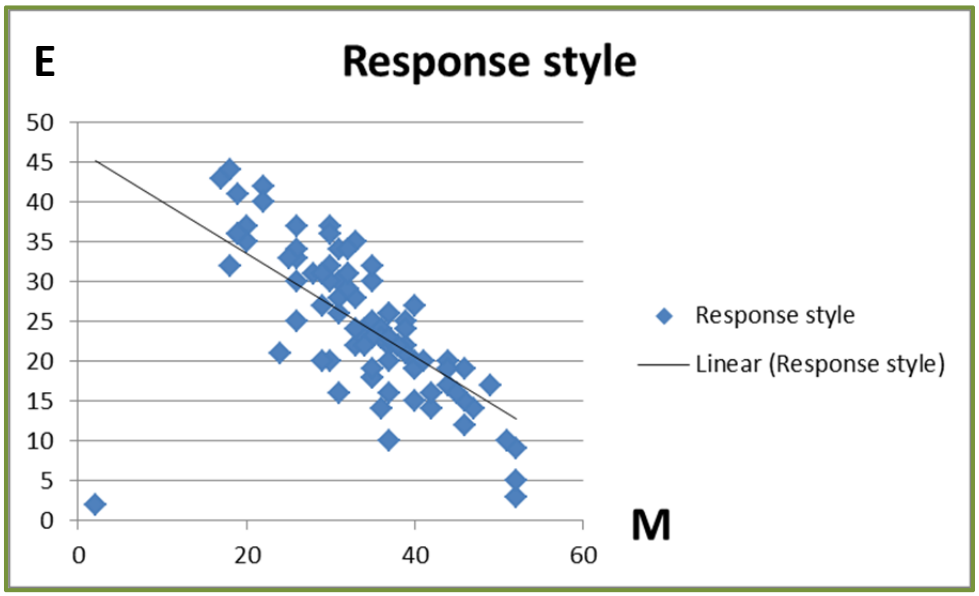

Fig 9. Scatter plot showing negative correlation between MRS and ERS for all questionnaires considered together.

There are also some intriguing correlations between response styles and entropy of responses (**Table 45**).

Table 45. Correlations between response styles and entropy of responses, using Spearman's *rho.*

rho	AHQ-SE	TAS_SE	BB-SE	BAQ-SE	PANAS-SE	MMCL-SE
AHQ-M	+ **					
AHQ-E	- *					
TAS-M		- n.s.				
TAS-E		+ **			+ **	
BB-M			- **		- *	
BB-E			+ **		+ *	
BAQ-M				+ n.s.		
BAQ-E				- n.s.		
PANAS-M					- n.s.	
PANAS-E		+ **			+ **	-*
MMCL-M						- **
MMCL-E						+ **

** = correlation significant at the 0.01 level; * = significant at the 0.05 level;
+ = positive correlation; - = negative correlation.

For four scales, correlation is positive between Shannon entropy and Extreme response style, but for one it is negative (and for another negative but nonsignificant). In contrast, for one scale correlation is positive between entropy and Medium response style (and positive for a further scale), but for two it is negative (and negative but nonsignificant for a further two scales).

It makes intuitive sense that a more Extreme response style (i.e., with more variance in responses) is likely to correspond to higher values of Shannon entropy; cf above, p 18). SE has been used to assess informativity of different questionnaires [Kim et al. 2012]. It is maximal when each of the questionnaire responses occurs with same probability (so that there is a lot of uncertainty) and is minimal when all responses are the same (for instance, if all are missing, so scored 'na'), with no

resulting uncertainty. It thus differs from standard deviation (or variance), but for unimodal distributions they share a monotonic relationship [Kondamudi 2015].

This was investigated further by assessing the degree of association between the standard deviation (SD), or the square root of the variance, of each individual's scores on a particular questionnaire with the SE of those same scores, and also between RS and standard deviations (**Table 46**).

Table 46. Associations between Standard deviation (SD), SE, MRS and ERS for each questionnaire.

rho	AHQ-SD	TAS_SD	BIS/BAS-SD	BAQ-SD	PANAS-SD	MMCL-SD
AHQ-SE	0.723					
AHQ-M	0.383					
AHQ-E	n.s.					
TAS-SE		0.773				
TAS-M		-0.435	- *			
TAS-E		0.717			*	
BIS/BAS-SE			0.772			
BIS/BAS-M			-0.582	- *	- *	
BIS/BAS-E		*	0.614	**	*	
BAQ-SE				0.618		
BAQ-M		- **		-0.473	- **	
BAQ-E		**		0.467	**	
PANAS-SE		**		*	0.611	
PANAS-M					-0.777	
PANAS-E		*		*	0.865	
MMCL-SE						0.799
MMCL-M	*					-0.847
MMCL-E				- *		0.928

All entries significant at the 0.001 level unless otherwise indicated.

For some measures such as BIS/BAS and BAQ, scatter plots showed a recognisable logarithmic distribution, but this was less marked for other measures such as I-PANAS-SF or TAS (**Figure 10**). This difference was independent of the amount of missing data per questionnaire.

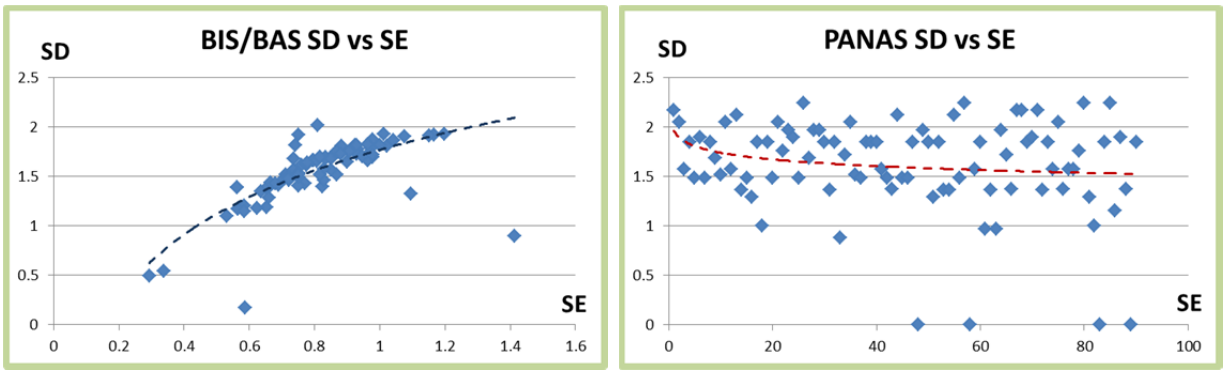

Fig 10. Scatter plots of SE vs SD for two different scales, BIS/BAS (*rho* = 0.772) and I-PANAS-SF (*rho* = 0.611).

The scatter plots reveal that SE and SD are not measuring the same construct, even if results are well correlated using Spearman's *rho*.

There were also a number of correlations between SD and RS or ES of *dIfferent* scales, but only six of these were significant at the 0.01 level (**Table 46**).

Hypothesis 13.1.2. Individuals may complete all questionnaires in a similar way, as assessed with Shannon entropy

Correlations between Shannon entropies (whether normalised or not) for the different scales were generally low and nonsignificant, except for that between TAS and PANAS (*rho* = 0.316, p = 0.002).

In contrast, there were significant differences among them, as indicated by the Friedman test: p<0.001 (with Chi-square = 108.401, df = 5, *N* = 90). This difference is in part due to the very low values of SE for the MMCL, for which there was so much missing data (22 questionnaires unanswered because of printing issues), cases with missing data resulting in a SE of zero. However, even without the MMCL, the Friedman test shows a significant difference among the remaining scales: p < 0.001 (Chi-square 100.53, df = 4, *N* = 90). Pairwise differences using the Wilcoxon test are shown in **Table 47**.

Table 47. Differences in normalised SE, pairwise, between the different questionnaires, showing *Z*, p values and effect size *r*.

nSE diffs	AHQ	TAS	BIS/BAS	BAQ	I-PANAS-SF	MMCL
AHQ		n.s.	*Z* = -6.118 p<0.001 *r* = 0.64	*Z* = -1.966 p=0.049 *r* = 0.21	*Z* = -2.273 p=0.023 *r* = 0.24	*Z* = -4.877 p<0.001 *r* = 0.51
TAS			*Z* = -7.603 p<0.001 *r* = 0.80	n.s.	n.s.	*Z* = -5.052 p<0.001 *r* = 0.53
BIS/BAS				*Z* = -7.691 p<0.001 *r* = 0.81	*Z* = -6.798 p<0.001 *r* = 0.72	n.s.
BAQ					n.s.	*Z* = -4.275 p<0.001 *r* = 0.45
I-PANAS-SF						*Z* = -4.095 p<0.001 *r* = 0.43
MMCL						

Thus correlations for SE between the questionnaires are small and infrequent, whereas differences are more likely (except for TAS). There does not seem to be a 'Shannon entropy response style' across different questionnaires. In confirmation of this, cluster analysis based on SE shows close links among only two or three of the primary scales used, with the MMCL in a cluster on its own.

Discussion

This study was undertaken as a continuation of our prior research into methods of measuring mood, exploring what personality and attitude factors might affect the outcomes of acupuncture treatment, and investigating personal styles in how people respond to questionnaires. Several of the present findings look sufficiently interesting and/or statistically significant to merit further investigation in future research.

Primary measures – Questionnaires

Analysis of some of our questionnaire data provides results that might be expected. Thus the AHQ suggests that there are two broad categories of condition – Physical and Functional versus Emotional/mental – for which people seek help from acupuncture (pp 8-9), and the TAS results could indicate that among traditional acupuncture students and practitioners, attitude to technology may be less positive than in the general population (p 9).

Other results are more interesting. The BIS/BAS behavioural inhibition scores, for instance (p 9), indicate that those attending EA courses may exhibit more rather than less exploratory behaviour, but without a particular drive for reward. This perhaps offers an interesting insight into the psychology of acupuncturists. The I-PANAS-SF scores suggest that acupuncturists may show more positive than negative affect, even though they may be less positive than the population at large (p 11).

The JTT results were mixed. They confirmed our earlier finding that acupuncturists may be more Feeling than Thinking, as well as perhaps more Intuitive than Sensing and more Perceiving than Judging (pp 11-12).

Some results are counterintuitive. The BAQ scores suggest that acupuncturists are *less* physically aware of how they feel than the population as a whole (p 10). This is a somewhat surprising result, given that acupuncturists are, in some ways, trained to be aware of their own and their patients' bodies and responses. It is, of course, possible that the BAQ does not reflect this kind of acupunctural awareness.

It is interesting that no particular stimulation parameters stood out for either intensity or pleasantness (p 13). Although HF stimulation was more often reported as most intense and least pleasant, LF as most pleasant and least intense, this may reflect the lecturer's declared bias.

Secondary measures –How questionnaires were completed

As for the secondary measures, the finding that for TAS, BIS/BAS, I-PANAS-SF and MMCL, there were more MRS than ERS scores, and for AHQ and BAQ, more ERS scores (p 14), should alert researchers to their potential benefits and/or shortcomings when using them in the future.

It may be worthy of note that across several questionnaires in this study women used more of an ERS and men more a MRS, and that those in the upper quartile of ERS scorers tended to be younger, and in the upper quartile of MRS scorers younger – although there is lack of agreement about such trends in the RS literature (p 14).

From their MMCL scores, an attempt was made to define who were 'good responders', 'non-responders' and 'poor responders' (pp 15-16), and it was clear that some of the MMCL items were

less responsive to acupuncture than others, with Positive and Relaxed being the most responsive, and Confused and Dispirited the least responsive of all ten items (p 15-16).

SE and nSE as measures of variability of responses varied for the different questionnaires. There was a tendency for older respondents to demonstrate lower SE when completing some questionnaires. Given that complexity and variability of many disparate physiological measures tend to decrease with age, it would be of interest to investigate in a larger study whether questionnaire SE follows a similar pattern. Furthermore, for all scales other than MMCL, SE was lower for men than for women – in keeping with the nonsignificant finding that MRS tends to be higher in men, and ERS in women (p 18).

The multiple mood change Likert instrument (MMCL)

Internal consistency of the MMCL was good when 'desirable' and 'undesirable' moods were analysed separately (p 21). Thus hierarchical cluster analysis permits a two-cluster solution – but also a three-cluster solution in which Alert and Energetic comprise the third cluster (p 21). Using the instrument on larger samples could help decide if the two- or three-cluster solution is more generally applicable. Convergent validity (using Spearman's *rho*) appears to be supported for the two (or three) clusters identified, with Tiredness rather out on a limb on its own (p 22). Thus **Hypothesis 1** (that the MMCL shows acceptable validity and internal consistency) is supported, in part, by the evidence.

Sensitivity of the MMCL to mood changes is somewhat greater than that of the NRS-M and $BRUMS_{24}$ (pp 23-25, contradicting **Hypothesis 2** (that the MMCL is less sensitive to change in mood than pre-to-post differences in NRS-M or $BRUMS_{24}$ scores), although there was a much higher percentage of no change scores for the former than the two previously used scales (pp 25-26). **Hypothesis 3** (that some respondents may consistently report larger or smaller mood changes than the majority) was only partially borne out, with only three in the upper quartile for mood change scores, and 20 in the lower quartile (p 26). Skewness of responses is less for the MMCL than for $BRUMS_{24}$, suggesting that it may be a more dependable measure (p 27). However, there is little difference between the three mood scales for test discrimination, $delta_G$, with the MMCL discriminating better between individuals on the 'desirable' than the 'undesirable' moods (p 29).

A central question, and the basis of ongoing research, is whether mood changes in response to treatment are affected by personality type or attitude (**Hypothesis 4**). Of some 50 possible associations between mood and type/attitude, only seven were significant, and none of these remained so after a simple Bonferroni correction (dividing p values by 50) (p 30). Considerable further research will clearly be needed, using additional outcome measures, to elucidate this problem.

Other primary measures of personality and attitude

Hypothesis 5, that benefit from prior acupuncture as assessed using the AHQ may be reflected in mood changes found during the teaching session, was supported only be three of a possible fifty tests (p 30), and the statistical significance of these three was low (and did not remain so after another simple Bonferroni correction). It would probably take a very large sample to investigate this further (a sample size calculation was not conducted).

Hypotheses 6 to **8**, on interactions between the primary measures and the EA stimulation parameters described as most or least pleasant and/or intense, were not supported by the evidence (pp 31-32). In contrast, **Hypothesis 9** (whether acupuncture is reported as beneficial in the past may depend on personality type or attitude) would be worth investigating in further research using the (positive) I-PANAS-SF scale. Results here suggest that those with a more positive view of life may well report benefits from past acupuncture treatment (p 32).

Hypothesis 10 (that there may be differences in personality type and/or attitude among training institutions) was perhaps supported for the BIS/BAS BAS-Drive subscale (pp 33-34). It might be of interest to know how recruitment is conducted at the different institutions, or whether those with more 'Drive' are attracted, for instance, to a University in central London, and those with lower 'Drive' to a college in a city in the North of England.

Hypothesis 11 (on the incidence of missing data) could not really be investigated properly, because of printing issues at one of the training institutions (p 7). In contrast, investigating **Hypothesis 12** (on other significant interactions between primary and secondary measures) led to several intriguing results. Of these, the following may merit further investigation:

- greater bodily awareness may be associated with Feeling less depressed or Dispirited after treatment (p 34).
- A more positive attitude to life may be associated with greater relaxation following treatment (p 34).
- BAS-Drive may be greater in Extraverts than Introverts (p 35). However (p 36), it is not possible simply to replace the Jungian Introvert/Extravert dyad with BIS/BAS.
- Response style of the I-PANAS-SF show some highly significant correlations ($p < 0.001$) with AHQ subscales, BIS and BAS-Drive, and BAQ , and of the MMCL with the BAQ, all of which make intuitive sense and suggest that the I-PANAS-SF should be retained for use in further research (pp 37-38).
- JTT Perceiver and Judger types may exhibit different RS for some questionnaires, but a large sample would be required to confirm this (p 38).

Hypotheses 13 and **14**, on patterns of responsiveness, response style and SE, were partially confirmed. There was indeed a certain amount of agreement between response style for *some* of the questionnaires (p 39). Further analysis on the basis of RS showed the questionnaires fell into two distinct clusters (pp 39-40). There was also a tendency for ERS and SE to be positively correlated, although not for all questionnaires (p 41). Furthermore, the standard deviation of responses (SD) – a standardised measure of variance – also correlated strongly with MRS, ERS and SE, although clearly these all represent different constructs (p 42). Correlations were of course weaker between SD and RS or ES of *dIfferent* scales (pp 42-43). Correlations between values of SE for different scales were generally low and nonsignificant, whereas differences were sometimes highly significant ($p<0.001$). Thus there does not seem to be a 'Shannon entropy response style' across different questionnaires (p 43).

Limitations

For a number of reasons, the results of this pilot study should be treated with caution. Numbers in each group taught were small, and at one training institution (CICM) there were issues with questionnaire printing that meant that they had to be completed hurriedly rather than at leisure. For this reason, there may have been errors in completing the questionnaires. Thus the students at CICM were the least likely to complete questionnaires as requested, as against those attending the Danish CPD course, who were the most likely to do so. Part of this extreme difference could also be explained by differences in the length and time of day of the teaching sessions. The CICM sessions were only half-day, not full one- or (in Denmark) two-day sessions; also the two CICM sessions took place in the afternoon and evening, whereas the other courses were scheduled for morning or all-day attendance. Some respondents may also have found it difficult to take in the questionnaire wording or focus on subjective mood in a busy group setting.

Additionally, students may have been tired (e.g., at the end of a one- or two-day course) or bored (e.g., if attendance on a course was compulsory), may have been uninterested in research, feeling that filling in questionnaires was a diversion from what they had come to learn, or may anyway consider mood changes as secondary and unimportant as a response to acupuncture treatment.

Moreover, it should be remembered that respondents' evaluation of change was *retrospective*, not prospective, and so may have been influenced by their state of mind post-treatment. Finally, issues with printing for students on one of the CICM courses led to large amounts of missing data.

Results of comparisons between the MMCL (N = 68) and $BRUMS_{24}$ or the NRS-M (N =158) can only be provisional. In addition, it is worth emphasising that the MMCL is probably more useful as a tool to assess *immediate* rather than longterm responses to treatment. A prospective comparison of the three measures on a larger single group of students or, indeed, patients should be undertaken to assess their relative usefulness for both immediate and longterm mood change assessment.

Conclusions and future directions

Despite problems with missing data and the type I errors that readily appear in exploratory research involving multiple comparisons, this small study does suggest some future avenues for investigation.

The multiple mood change Likert scale (MMCL) introduced here shows acceptable validity, internal consistency and sensitivity when compared to existing scales used to assess mood changes, such as the $BRUMS_{24}$ or our multiple NRS for mood (NRS-M). In our next study, we intend to compare these 'head-to-head', prospectively rather than retrospectively, and to assess the performance of the MMCL using ROC (receiver operating characteristics) analysis.

Disappointingly, interactions between the other primary measures of personality, attitude and helpfulness of prior acupuncture used in this study were not particularly revealing. In particular, the Jungian typology test (JTT) was less useful than expected. In contrast, however, affective style as assessed using the I-PANAS-SF strongly influenced reporting of benefits from prior treatment, a potentially important finding. A different (but overlapping) selection of personality scales will therefore be used in our next study of mood changes in response to treatment.

The secondary measures used – response style (RS), responsiveness and Shannon entropy (SE) – were all of interest, with the primary measures used falling neatly into two RS clusters, BIS/BAS and BAQ versus the rest, for example, although this was not confirmed by SE cluster analysis. There were strong correlations of questionnaire standard deviation (SD) with MRS, ERS and SE; further study would be helpful to elucidate whether SD or SE is a more useful measure.

Conflict of interests

None declared.

Acknowledgements

To the University of Hertfordshire for granting Ethics approval for this study; to the training institutions concerned for granting access to those attending electroacupuncture seminars for the purposes of the study; to the attendees who took the time to complete questionnaires that may for some have seemed to have been of no personal relevance whatsoever; to Tim Watson at the University of Hertfordshire; to Mike Cummings for his salutary comments on multiple comparisons; and of course to our spouses/partners, without whom none of this would have been possible at all.

References

Abu-Bader SH. (2010). Advanced and Multivariate Statistical Methods for Social Science Research with a Complete SPSS Guide. Chicago, IL: Lyceum Books.

Anon. (n.d.). Cronbach's Alpha. http://www.open.ac.uk/socialsciences/spsstutorial/files/tutorials/cronbachs-alpha.pdf [accessed 28.01.18].

Anon. (n.d.) Effect size of a binomial test and its relation to other measures of effect size. https://stats.stackexchange.com/questions/176847/effect-size-of-a-binomial-test-and-its-relation-to-other-measures-of-effect-size [accessed 09.12.17].

Anon. (2015). Tsallis and Rényi Normalized Entropy. https://stats.stackexchange.com/questions/134748/tsallis-and-r%c3%a9nyi-normalized-entropy [accessed 02.02.18].

Briggs Myers I, Myers PB. (1995). *Gifts Differing: Understanding personality type*. Palo Alto, CA: Davies-Black Publishing.

Carver CS, White TL. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of Personality and Social Psychology* 67(2): 319-333.

Glynn LM, Howland MA, Sandman CA, Davis EP, Phelan M, Baram TZ, Stern HS. (2018). Prenatal maternal mood patterns predict child temperament and adolescent mental health. *Journal of Affective Disorders* 228:83-90.

Handayani PW, Hidayanto AN, Pinem AA, Sandhyaduhita PI, Budi I. (2017). Hospital information system user acceptance factors: User group perspectives. *Informatics for Health and Social Care* 1-24.

Hankins M. (2008). How discriminating are discriminative instruments? *Health and Quality of Life Outcomes* 6:36.

Jung CG, Haynes HG. (1971). *Psychological types*. Collected works of CG Jung, 6. London: Routledge.

Kim SH, Kim HJ, Lee SI, Jo MW. (2012). Comparing the psychometric properties of the EQ-5D-3L and EQ-5D-5L in cancer patients in Korea. *Quality of Life Research* 21(6):1065-73.

Kondamudi S. (2015). Which is the best measure of uncertainty, variance or entropy, or are they both equivalent? [https://www.quora.com/Which-is-the-best-measure-of-uncertainty-variance-or-entropy-or-are-they-both-equivalent [accessed 22.01.18]

Manor B, Lipsitz LA. (2013). Physiologic complexity and aging: implications for physical function and rehabilitation. *Progress in Neuro-psychopharmacology and Biological Psychiatry* 45:287-293.

Mayor D, Steffert T. (2016). Measuring mood – relative sensitivity of numerical rating and Likert scales in the context of teaching electroacupuncture. Initial findings and the influence of response style on results. Poster, ARRC Symposium, London, 2016 [http://www.qeeg.co.uk/electroacupuncture/EAmood.htm].

Pincus SM, Schmidt PJ, Palladino-Negro P, Rubinow DR. (2008). Differentiation of women with premenstrual dysphoric disorder, recurrent brief depression, and healthy controls by daily mood rating dynamics. *Journal of Psychiatric Research* 42(5):337-347.

Rosen LD, Whaling K, Carrier LM, Cheever NA, Rokkum J. (2013). The Media and Technology Usage and Attitudes Scale: an empirical investigation. *Computers in Human Behavior* 29(6): 2501-2511.

Shields SA, Mallory ME, Simon A. (1989). The Body Awareness Questionnaire: reliability and validity. *Journal of Personality Assessment* 53(4): 802-815.

Shannon CE. (1948). A mathematical theory of communication. *Bell System Technological Journal* 27: 379-423, 623-656.

Steffert T, Mayor D. (2017). Mood changes in response to electroacupuncture treatment in a classroom situation. Personality type, emotional intelligence and prior acupuncture experience, with an exploration of Shannon entropy, response style and graphology variables. Poster, ARRC Symposium, London, 2017 [http://www.qeeg.co.uk/electroacupuncture/moodchange.htm].

Tastle WJ, Wierman MJ. (2006). An information theoretic measure for the evaluation of ordinal scale data. *Behavior Research Methods* 38(3):487-494.

Thompson ER. (2007). Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology* 38(2): 227-242.

Van Vaerenbergh Y, Thomas TD. (2013). Response styles in survey research: a literature review of antecedents, consequences, and remedies. *International Journal of Public Opinion Research* 25(2):195-217.

Young ME, Wasserman EA. (2001). Entropy and variability discrimination. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 27(1):278-93.

APPENDIX A. Calculating Shannon entropy in Excel

In this and our previous study on the Shannon entropy of questionnaire responses, H(X) or SE, we used an online calculator available at http://www.shannonentropy.netmark.pl/calculate/. However, this requires careful data entry to avoid data from Excel spreadsheets being entered with a space between each data point, resulting in erroneous values for H(X). Furthermore, each individual's responses has to be entered separately, which is time consuming and increases the likelihood of errors.

Other Shannon entropy calculators located online may give different results, and calculators for the ecological diversity 'Shannon index' lead to different results again.

It was therefore decided in any future studies to go back to first principles [Shannon 1948] and use standard Excel functions to calculate Shannon entropy for questionnaires with a small number of possible responses, such as Likert or numerical rating scales, but not analogue scales *per se*. This procedure can be used sequentially on data from multiple respondents without having to re-enter it, and so is considerably faster than the other methods located online. For those familiar with the programming language MATLAB , code for SE is readily available, and iterative calculation will be faster still.

Method

If your data are not already in an Excel spreadsheet, enter them so that each data string (e.g., for a Likert-type questionnaire, the numerical responses for each person's questionnaire responses) are entered in rows, starting with Row 1, Column Q (and to the right until the last item is reached for that person's responses). Enter all your respondents' data, but *not* in consecutive rows. Instead, leave 5 rows empty between each person's responses. So if the first person's responses are in Row 1, the next would be in Row 7, the one after in row 13, and so forth.

Check the data for any empty cells or non-numerical entries, and replace empty cells or those with non-numerical entries with a standard string of letters (such as 'na').

Next, in Row 1, starting in Column A, create a row of cells that covers all possible questionnaire responses, e.g. 1 to 7. The aim is to count how many times each response is given, using the Excel function '=Countif(range, response)'. Using the data from all the questionnaires in this study as an example, this would result in 11 cells being filled, from -2 to +7, with 'na' as an extra cell. If the range of cells occupied by the questionnaire responses is P1 to P10, say, Row 1 (A1 to K1) would look like:

=countif(Q1:Q10,-2)	=countif(Q1:Q10,-1)	…	…	…	…	=countif(Q1:Q10,7)	=countif(Q1:Q10,"na")

Then in Cell M1 count how many columns there are in the range (i.e., how many questions there are in the measure, *N*), using '=Columns(range)' if necessary.

In another row (Row 2, A2 to K2), divide the amounts in Row 1 by the number of columns in the questionnaire range from cell M1 (so A2=A1/$M$1, B2=B1/$M$1, etc.). As a result, Row 2 will probably contain some cells that just contain a zero.

In Row 3 (A3 to K3), use a function that will turn all these zeros into a number that is NOT one used as a questionnaire response, such as '10' For example, if Row 2 runs from cell A2 to K2, use =IF(A2=0,10,A2), =IF(B2=0,10,B2), … to =IF(K2=0,10,K2):

=IF(A2=0,10,A2)	=IF(B2=0,10,B2)	…	…	…	…	…	=IF(K2=0,10,K2)

In Row 4 (A4 to K4), derive the logarithm to the base 2 of the data in the row above by entering =LOG(A3,2), =LOG(B3,2) … to =LOG(K3,2):

=LOG(A3,2)	=LOG(B3,2)	…	…	…	…	…	=LOG(K3,2)

In row 5 (A5 to K5), enter the product of the cells in the same column in Rows 3 and 4, i.e.:

=A3*A4	=B3*B4	…	…	…	…	…	=K3*K4

Then in Cell M2, sum the cells in Row 5, i.e. =SUM(A5:K5), and in Cell M3 enter the result of =COUNTIF(A5:K5, 10*LOG(10,2)). In Cell M4, enter =10*M3*LOG(10,2).

Finally, in Cell N4 deduct the sum of the cells in row 5 from the amount in M4, so N4=M4-M2.

This is the Shannon entropy for your questionnaire string.

For normalised Shannon entropy, in cell O4 divide the Shannon entropy by $Log_2 N$, i.e. O4=N4/log(M1,2).

Copy cells A1 to O5, and position them so that Cell A1 is copied to Cell A7, level with the next row of questionnaire responses, Row 7.

This should provide you with values of SE and nSE for the next person's questionnaire responses, in cells N10 and O10, respectively.

Continue copying until you have calculated SE and nSE for all your respondents.

In Column P, label or number the rows containing the SE and nSE results according to the original labels for the respondents, to facilitate sorting the results afterwards. Copy Columns N to P, starting with the first row of results, and then paste the *values* where you want to put them, sorting them in order by using the numbers or labels in Column P.

APPENDIX B. Hand-out of primary measures used in this study

Class exercise on electroacupuncture (EA)

Part 1. PRELIMINARY QUESTIONNAIRES FOR YOU TO COMPLETE

Context: Various learning tools are being piloted with your group. This is part of a larger study to explore the effects of different frequencies of electroacupuncture (EA) or TENS at acupuncture points (TEAS) on the electrical activity of the brain (EEG).
Objectives: One objective of this exercise is to explore which individual characteristics and attitudes may affect your experience of electroacupuncture.
Your role: You will be asked to complete some online and paper questionnaires before and during the EA teaching session. No personal information about you will be recorded, and you do not need to take part if you do not wish to.
If you have any questions about the questionnaires or the study and your role in it, you can contact me (David Mayor) at davidmayor@welwynacupuncture.co.uk or 01707 320782.

Please complete the following questionnaires **before** your electroacupuncture teaching session (and when you have done so **bring them with you** to the session).

1.1. TAS (items from the Technology Attitudes Scale – less than 2 minutes)

Instructions

Please answer the following questions about your attitudes to technology. For each question, select a number from 1 to 5 that best describes how the statement describes you and place the number in the box to the right of the question.

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1	2	3	4	5

1. I think it is important to keep up with the latest trends in technology. ☐
2. I am dependent on my technology. ☐
3. Technology will provide solutions to many of our problems. ☐
4. With technology anything is possible. ☐
5. I feel that I get more accomplished because of technology. ☐
6. New technology makes people waste too much time. ☐
7. New technology makes life more complicated. ☐
8. New technology makes people more isolated. ☐

[Rosen LD, Whaling K, Carrier LM et al. (2013). The Media and Technology Usage and Attitudes Scale: An empirical investigation. *Computers in Human Behavior*, 29(6):2501-2511]

** Please turn over ...

1.2. BIS/BAS (Behavioural Inhibition and Activation Scales – less than 5 minutes)

Instructions

Please read all these instructions carefully before beginning to complete the questionnaire.

The questionnaire contains 24 statements. Read each statement carefully.

Each statement is one that a person may either agree with or disagree with.

For each statement, indicate **how much you agree or disagree** with what the statement says. Please respond to all the statements; **do not leave any blank**. You can only choose one response to each statement.

Please be as accurate and honest as you can be. Respond to each statement as if it were the only item. That is, don't worry about being 'consistent' in your responses. There are no 'correct' or 'incorrect' answers. Do not spend too much time on one statement and answer according to your own feelings, rather than how you think 'most people' would feel.

Use the following scale to record your answers (choose only one response to each statement), selecting a number from 1 to 4 and placing the number in the box to the right of the statement.

Very true for me	**Somewhat true for me**	**Somewhat false for me**	**Very false for me**
1	**2**	**3**	**4**

1. A person's family is the most important thing in life. ☐
2. Even if something bad is about to happen to me, I rarely experience fear or nervousness. ☐
3. I go out of my way to get things I want. ☐
4. When I'm doing well at something I love to keep at it. ☐
5. I'm always willing to try something new if I think it will be fun. ☐
6. How I dress is important to me. ☐
7. When I get something I want, I feel excited and energised. ☐
8. Criticism or scolding hurts me quite a bit. ☐
9. When I want something I usually go all-out to get it. ☐
10. I will often do things for no other reason than that they might be fun. ☐
11. It's hard for me to find the time to do things such as get a haircut. ☐
12. If I see a chance to get something I want I move on it right away. ☐
13. I feel pretty worried or upset when I think or know somebody is angry at me. ☐
14. When I see an opportunity for something I like I get excited right away. ☐
15. I often act on the spur of the moment. ☐
16. If I think something unpleasant is going to happen I usually get pretty "worked up." ☐
17. I often wonder why people act the way they do. ☐
18. When good things happen to me, it affects me strongly. ☐
19. I feel worried when I think I have done poorly at something important. ☐
20. I crave excitement and new sensations. ☐
21. When I go after something I use a "no holds barred" approach. ☐

** Please turn over ...

22. I have very few fears compared to my friends. ☐
23. It would excite me to win a contest. ☐
24. I worry about making mistakes. ☐

[Carver CS, White TL (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67(2), 319-333]

1.3. BAQ (Body Awareness Questionnaire – less than 4 minutes)

Instructions

Listed below are a number of statements regarding your sensitivity to normal, non-emotive body processes. For each statement, select a number from 1 to 7 that best describes how the statement describes you and place the number in the box to the right of the statement.

Not at all true of me						**Very true of me**
1	**2**	**3**	**4**	**5**	**6**	**7**

1. I notice differences in the way my body reacts to various foods. ☐
2. I can always tell when I bump myself whether or not it will become a bruise. ☐
3. I always know when I've exerted myself to the point where I'll be sore the next day. ☐
4. I am always aware of changes in my energy level when I eat certain foods. ☐
5. I know in advance when I'm getting the flu. ☐
6. I know I'm running a fever without taking my temperature. ☐
7. I can distinguish between tiredness because of hunger and tiredness because of lack of sleep. ☐
8. I can accurately predict what time of day lack of sleep will catch up with me. ☐
9. I am aware of a cycle in my activity level throughout the day. ☐
10. I ***don't*** notice seasonal rhythms and cycles in the way my body functions. ☐
11. As soon as I wake up in the morning, I know how much energy I'll have during the day. ☐
12. I can tell when I go to bed how well I will sleep that night. ☐
13. I notice distinct body reactions when I am fatigued. ☐
14. I notice specific body responses to changes in the weather. ☐
15. I can predict how much sleep I will need at night in order to wake up refreshed. ☐
16. When my exercise habits change, I can predict very accurately how that will affect my energy level. ☐
17. There seems to be a "best" time for me to go to sleep at night. ☐
18. I notice specific bodily reactions to being over-hungry. ☐

[Shields SA, Mallory ME, Simon A (1989). The Body Awareness Questionnaire: reliability and validity. *Journal of Personality Assessment*, 53(4), 802-815]

** Please turn over ...

1.4. PANAS-SF (Positive and Negative Affect Schedule Short Form – less than 2 minutes)

Instructions

Thinking about yourself and how you normally feel, select a number from 1 to 5 that best describes you and place the number in the box to the right of the question.

Never				**Always**
1	**2**	**3**	**4**	**5**

To what extent do you generally feel:

1.	Upset	☐
2.	Hostile	☐
3.	Alert	☐
4.	Ashamed	☐
5.	Inspired	☐
6.	Nervous	☐
7.	Determined	☐
8.	Attentive	☐
9.	Afraid	☐
10.	Active	☐

[Thompson ER (2007). Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology*, 38(2): 227-242]

1.5. Myer Briggs personality type

If you do not already know your Myers Briggs type, please take the free online 'Jung Typology Test' at **http://www.humanmetrics.com**. It does not take long to do. Then circle your type in the Table below. (Note: Your Keirsey 'temperament' and Jungian 'type' may be slightly different.)

If you already know your Myers Briggs personality type, please circle below:

ISTJ	Introverted Sensing Thinking Judging (Introverted Sensing with Extraverted Thinking)	ESTP	Extraverted Sensing Thinking Perceiving (Extraverted Sensing with Introverted Thinking)
ISFJ	Introverted Sensing Feeling Judging (Introverted Sensing with Extraverted Feeling)	ESFP	Extraverted Sensing Feeling Perceiving (Extraverted Sensing with Introverted Feeling)
INFJ	Introverted iNtuitive Feeling Judging (Introverted Intuition with Extraverted Feeling)	ENFP	Extraverted iNtuitive Feeling Perceiving (Extraverted Intuition with Introverted Feeling)
INTJ	Introverted iNtuitive Thinking Judging (Introverted Intuition with Extraverted Thinking)	ENTP	Extraverted iNtuitive Thinking Perceiving (Extraverted Intuition with Introverted Thinking)
ISTP	Introverted Sensing Thinking Perceiving (Introverted Thinking with Extraverted Sensing)	ESTJ	Extraverted Sensing Thinking Judging (Extraverted Thinking with Introverted Sensing)
ISFP	Introverted Sensing Feeling Perceiving (Introverted Feeling with Extraverted Sensing)	ESFJ	Extraverted Sensing Feeling Judging (Extraverted Feeling with Introverted Sensing)
INFP	Introverted iNtuitive Feeling Perceiving (Introverted Feeling with Extraverted Intuition)	ENFJ	Extraverted iNtuitive Feeling Judging (Extraverted Feeling with Introverted Intuition)
INTP	Introverted iNtuitive Thinking Perceiving (Introverted Thinking with Extraverted Intuition)	ENTJ	Extraverted iNtuitive Thinking Judging (Extraverted Thinking with Introverted Intuition)

** Please turn over ...

Please enter your year of birth and gender in the boxes below:

Year of birth ☐

Gender (F/M) ☐

1.7. Have you had **acupuncture treatment** in the past, or are you having it currently?

Tick a box:

Yes ☐ No ☐

1.7.1. If you answered 'yes', have you found acupuncture generally helpful for yourself? Please circle the number you consider appropriate in the box below.

1 . . . . . 2 . . . . . 3 . . . . . 4 . . . . . 5 . . . . . 6 7

Not helpful at all — Extremely helpful

1.7.2. In particular, have you found it helpful for:

	(not at all)						(extremely)
Purely physical or biomedical problems	1	2	3	4	5	6	7
'Functional' disorders or medically unexplained symptoms	1	2	3	4	5	6	7
Mental/emotional difficulties	1	2	3	4	5	6	7
Lifestyle or other advice given	1	2	3	4	5	6	7

Please confirm that you are willing that the questionnaires you complete for this teaching session can be used for the lecturer's stated research purposes:

Tick box:

☐

Password:

NB: Please remember to bring these questionnaires to the electroacupuncture lecture!

** Please turn over ...

Part 2. PRACTICAL EXERCISE

Instructions

In groups of three – 'practitioner', 'patient' and 'observer' – take turns to give and receive treatment. Each of you should experience no more than 5 minutes of ALL the following (in any order), i.e. for a total of 20 minutes or less:

- 'low frequency' stimulation (2-4 Hz)
- 'midrange' stimulation (around 10-15 Hz)
- 'high frequency' stimulation (80-100 Hz)
- 'dense-disperse' or 'modulated' stimulation (e.g. alternating low and high frequencies).

Please make sure you experience **ALL FOUR** treatments. Use the same acupuncture points for all four treatments.

The 'practitioner' sets the frequency (and pulse duration) first, and then the amplitude. The amplitude is determined by what the 'patient' experiences as 'strong but comfortable'.

2.1. To be completed by the 'OBSERVER' – on the patient's copy of this hand-out

During treatment, record the following where possible:

2.1.1. Device name (e.g. AWQ104L, or ES-160): ……..

2.1.2. Parameters:

Order	**Points used**	**Frequency/mode**	**Amplitude setting**	**Pulse duration**	**EA duration**
example	*LI4/LI5**	*4 Hz (LF)*	*7dial units*	*300 µs*	*4 mins*
EA1					
EA2					
EA3					
EA4					

* Note that at CICM only off-meridian points should be used.

C. Any other observations:

** Please turn over …

2.2. To be completed by or with the 'PATIENT' after their treatment – on the patient's copy of this hand-out

Below you (the 'patient') will find two sets of questions. Please answer them according to your own feelings, as accurately and truthfully as you can, rather than how you think you 'should' feel.

2.2.1. Compared to how you felt just before the treatment you received today, do you now feel:

Feeling	**Considerably less …**	**Somewhat less …**	**Neither more nor less …**	**Somewhat more …**	**Considerably more …**	**Don't know/ not sure**
Alert	– 2	– 1	0	1	2	?
Anxious	– 2	– 1	0	1	2	?
Confused	– 2	– 1	0	1	2	?
Dispirited	– 2	– 1	0	1	2	?
Energetic	– 2	– 1	0	1	2	?
Negative	– 2	– 1	0	1	2	?
Pain	– 2	– 1	0	1	2	?
Positive	– 2	– 1	0	1	2	?
Relaxed	– 2	– 1	0	1	2	?
Tired	– 2	– 1	0	1	2	?

2.2.2. Which frequency/mode parameters did you find the most/least intense and pleasant? Please tick the four boxes below which you consider reflect most accurately what you felt.

frequency/mode	most intense	least intense	most pleasant	least pleasant
Low frequency				
Midrange frequency				
High frequency				
Dense-disperse mode				
Don't know/ not sure				

Any other observations:

If you would like to take part in another research project on who responds well to acupuncture, you can email me (davidmayor@welwynacupuncture.co.uk) for more information.

If you do decide to take part, you would simply be asked to complete some more questionnaires about yourself online. This would take you around 45-90 minutes.