Changes in finger temperature and blood flow in response to different frequencies of transcutaneous electroacupuncture at LI4 (*hegu*). Interim analysis and 'real life' methodological issues: many factors, missing data and a multiplicity of measures

© David Mayor[1], Tony Steffert[2] and Ronakben Bhavsar[3]

1. School of Health and Social Work, University of Hertfordshire; 2. Music Computing Lab, Centre for Research in Computing, Open University; 3. School of Computer Science, University of Hertfordshire

CONTENTS

ABSTRACT

Background. Finger blood flow and temperature are often used as indices of autonomic (sympathetic) function. Transcutaneous electroacupuncture (TEA) is an increasingly used variant of electroacupuncture (EA). This is the fifth in a series of conference posters from a study investigating the effects of EA and TEA on heart/pulse rate variability (HRV/PRV), the electroencephalograph (EEG), and now blood flow.

Objectives. To assess how treatment factors – particularly TEA stimulation frequency (Hz) – contribute to changes in local blood flow, skin temperature and pulse transit time (PTT), and to explore associations between the outcome measures used.

Methods. In this pilot study, participants (*N*=17) each attended for a single session consisting of 10 consecutive 5-minute 'slots'. In the second, fifth and eighth slots, TEA was applied bilaterally at LI4 (*hegu*) at three different frequencies (2.5 Hz, 10 Hz and 80 Hz), in counterbalanced order (all with 256 μs pulse duration and at a 'strong but comfortable' intensity). Using finger photoplethysmography, with a thermistor on the same finger, the blood volume pulse (BVP) and temperature were monitored throughout. Electrocardiograph (ECG) signals were collected from wrist electrodes. Data on blood flow amplitude, pulse rate and transit time were derived from the BVP and ECG following standard procedures, including artefact processing. Mood was assessed at various time points using numerical rating scales and a short form of the Profile of Mood States questionnaire. Sample size estimates for the different measures and experimental factors were conducted as a basis for continuation of the study.

Results. TEA at 2.5 Hz consistently but not significantly resulted in greater fingertip blood flow, and 80 Hz in longer PTT, than at the other two stimulation frequencies (frequency effects on temperature were inconsistent, small and not significant). For most participants, the association between skin blood flow and temperature was significant and positive, with both tending to peak together shortly after TEA. However, over the session both tended to decrease, although several measures of mood and PRV improved. In contrast, overall session change in PTT was small, but *group* medians peaked in slots 5 and 8.

Conclusions. 2.5 Hz TEA is likely to enhance local skin blood flow. However, even if this question is apparently simple, its investigation turned out to be complicated. As we found before when studying HRV, stimulation frequency may be a less important factor than others such as the presence of muscle twitch or participants' prior experience of related treatments. Thus in future our analysis will use multilevel modelling to take account of multiple factors and their interactions. In addition, we noted that although PTT may be helpful in assessing short-term changes in acupuncture-related research, this will only be so if high sampling rates are used. Further recruitment of participants is planned to consolidate our findings.

Background/Objectives

This is the fifth in a serious of pilot studies. In the earlier studies, we investigated the effects of different frequencies and locations of needle and transcutaneous electroacupuncture (TEA) on eyeblink rate and the electrical activity of the cerebral cortex using electroencephalography (EEG), and on heart rate and pulse rate variability (HRV, PRV) using forearm electrocardiography (ECG) and a fingertip blood volume pulse (BVP) sensor. The methods used and some of the outcomes obtained have been described in previous poster presentations [Mayor & Steffert 2013a, 2013b; Steffert & Mayor 2013, 2014]. Background information on EA and TEA can be found in standard textbooks [Johnson 2014; Mayor 2007].

The primary objective (1) of the present study was to assess whether cutaneous blood flow and temperature measured on one finger of the nondominant hand vary with different frequencies of TEA applied bilaterally at acupoint LI4 (*hegu*) and the ulnar border of the hand. Other objectives included (2) determination of the pulse transit time (PTT) between ECG and BVP pulses in order to assess its usefulness as a measure of short-term changes in response to TEA, and (3) to rank the effects on the outcome measures used of factors in the experimental situation other than stimulation frequency.

Peripheral cutaneous blood flow and temperature are both governed by the autonomic nervous system, with vasoconstriction and temperature decrease an indication of sympathetic dominance, and vasodilatation and temperature increase an indication that this is reduced [Kistler 1998]. HRV/PRV frequency domain indices are also governed by the autonomic nervous system, with vagal activity the major contributor to high frequency components, and its sympathetic modulation reflected in low frequency components [Malik 1996]. Increase in other HRV measures is often considered an objective measure of improved subjective relaxation [Steffert & 2014]. Therefore a further objective was to explore possible associations between outcome measures of blood flow, temperature, HRV/PRV and mood.

Methods

Design and recruitment

This is a single-arm, within-subject, repeated measures study, in which participants have each attended for a single session around 90-110 minutes long, during which TEA was applied at three different frequencies (2.5 Hz, 10 Hz and 80 Hz, all with 256 µs pulse duration), in counterbalanced order. Intensity has been set individually, in response to participants stating that they feel it as 'strong but comfortable'. Participants have therefore not been blinded to intervention.

Ethics approval for the study was granted by the University of Hertfordshire School of Health and Human Sciences Ethics Committee. Seventeen healthy individuals have taken part to date, recruited from among the students and staff at the University of Hertfordshire (N=15) using intranet announcements and, in some cases, a personal approach. In addition, two therapists working locally agreed to take part in the study. Six potential participants were excluded on health grounds (exclusion criteria may be found in **Appendix I** below).

Potential participants contacted the lead researcher (DM) by email and were then sent detailed information about what taking part in the study would involve. If they agreed in principle to

continue during an ensuing telephone conversation, they were sent individual codes for a brief online questionnaire, to be completed approximately a week before their experimental session.

The experimental session

The study has been conducted on the University premises, so far in one meeting room and one smaller staff office.

During their session, participants were given the opportunity to ask further questions and then signed a consent from. Next they were asked to complete a 5-minute questionnaire (Cohen's Perceived Stress Scale) [Cohen 1988] and a 5-minute composite numerical rating scale (Composite NRS, **Appendix II**) to assess general relaxation and mood over the past month, followed by two further short questionnaires (one being a short version of the Profile of Moods State, POMS-SF [Shacham 1983]) and another Composite NRS for current relaxation and mood (**Appendix II**). After this, seated in a comfortable chair with forearms supported, an EEG electrode cap was positioned (to gather data for a different research project) and other sensors were attached to arms and fingers to record heart rate, skin blood flow and temperature [**Fig 1**].

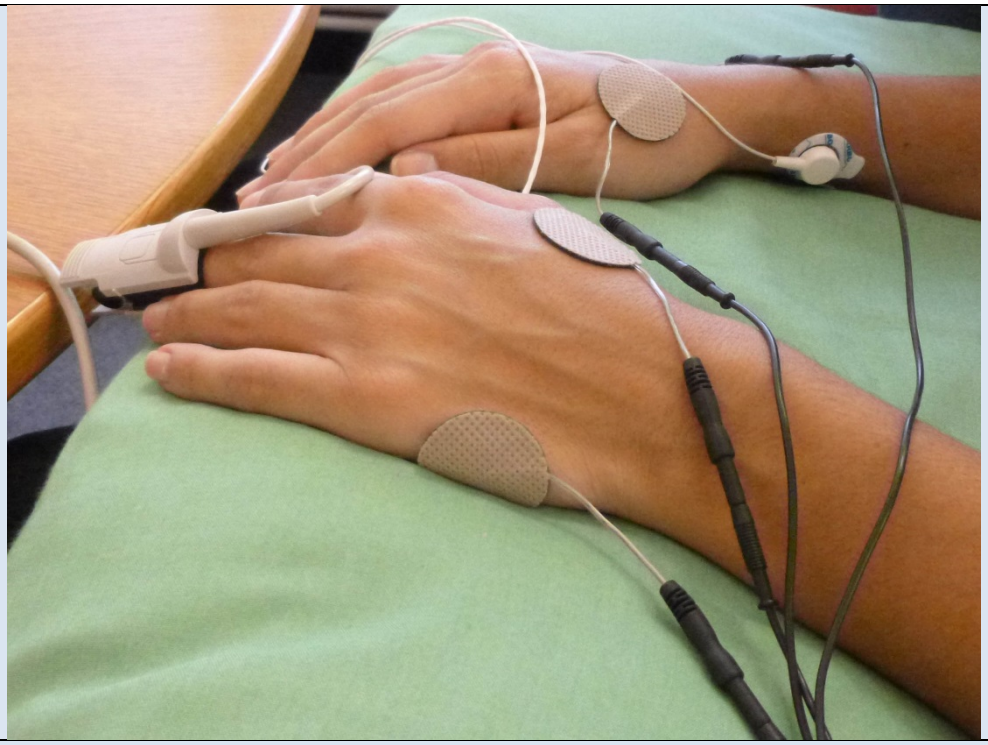


Fig 1. Sensors and electrodes in place, showing fingertip BVP sensor, one ECG electrode on right forearm, and TENS electrodes at LI4 and on the ulnar border of the hands. The temperature thermistor on the pad of the left middle finger is hidden by the fingertip sensor, and the ECG electrodes on the left forearm are also not visible.

ECG was collected using 3 pre-gelled disposable wrist electrodes (Kendall/Tyco ARBO H124SG, Henleys Medical Supplies, Welwyn Garden City) and an abdominal respiration rate sensor, and BPV using a fingertip plethysmograph. The temperature (thermistor) and BVP sensors were applied to the same finger (usually the middle, but sometimes the ring finger). All non-EEG signals were

recorded using the NeXus-10 physiological amplifier (Mind Media, Herten, Netherlands, also suppliers of the associated sensors). In addition, core body temperature was assessed using a standard oral thermometer with disposable probe covers (ADC, Hauppage, US). Room temperature was kept as constant as possible using a thermostatically controlled radiator (TRD4 1025E, De'Longhi. Treviso, Italy), monitored with a previously tested Gentle Temp 521 digital ear thermometer (Omron Healthcare, Milton Keynes). The Mind Media temperature sensor had previously been calibrated against a reference thermocouple device (MM 2020, TM Electronics, Worthing). Stimulation was carried out using a Class IIa acupuncture/TENS stimulator (Equinox E-T388, Liverpool) whose output had been checked for suitability – clean square waves, free of spikes, and stability of frequency output – in the University's Physiotherapy Research laboratory (two other devices were rejected for these reasons) [**Fig 2**]. Not being a constant current device, the stimulator output was only available in arbitrary 'dial units' rather than milliampères.

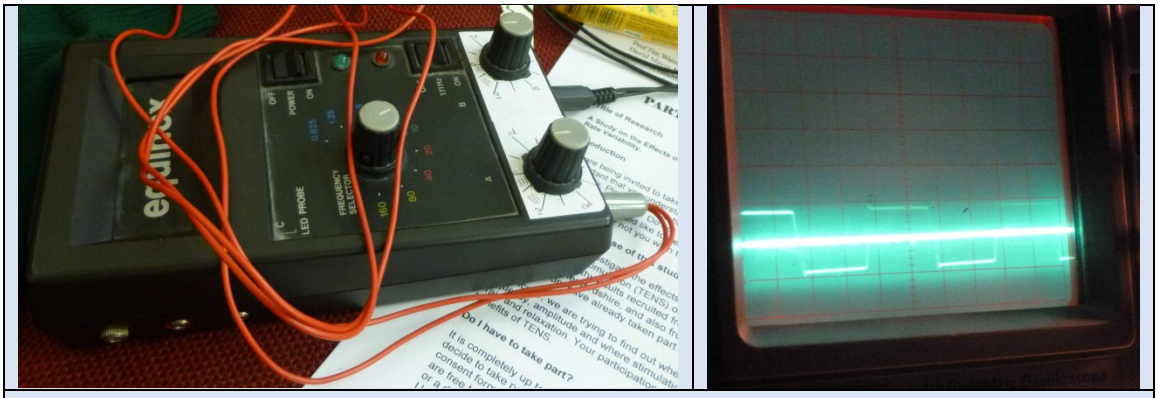


Fig 2. Left. The Equinox stimulator with label showing 'dial units'.
Right. Waveform from the device.

Following standard skin preparation with an alcohol wipe, self-adhesive TENS electrodes (Stimex, 32 mm diameter, Schwa-Medico, Ehringshausen, Germany) were attached to both hands and connected to the stimulator. Following an initial 5-minute baseline recording (with no stimulation), recording was continued and TEA applied for 5 minutes at one stimulation frequency and at a 'strong but comfortable' intensity selected by the participant. Once this stimulation level was accepted by the participant, a second 5-minute period of monitoring followed. At the end of this, stimulation level was decreased to zero and the unit switched off. Then followed a break of 10 minutes during which monitoring continued. Halfway through this break participants were asked to complete the same Composite NRS as before (with item order randomised) and were given the opportunity to move around to make themselves comfortable; oral temperature was taken again. Another 5 minutes of stimulation at a different frequency followed, again with a 10-minute break afterwards during which the measurement scale was completed and the oral temperature taken. This was repeated a third time, and then the various electrodes and other sensors were removed, and participants asked to complete the initial POMS-SF questionnaire again. They were also asked which of the three stimulation frequencies they liked or disliked most, and any other verbal feedback provided was recorded in writing. Following the session, participants completed a brief voluntary online follow-up questionnaire.

Data capture and measures used

Temperature and Respiration were recorded at 32 Hz (samples per second), ECG at 256 Hz and BVP at 128 Hz. Data was recorded and processed using BioTrace+ software (Mind Media, Herten, Netherlands). Further processing was conducted in Matlab, Excel 2010 and Kubios HRV v 2.1 2012 (Biosignal Analysis and Medical Imaging Group, University of Eastern Finland, Kuopio, FI), and statistical analysis using SPSS 20 (IBM, USA). The BioTrace+ output included the 'smoothed' BVP amplitude, calculated as the Root Mean Square (RMS) value of peak-to-peak amplitude (in µV) from the BVP sensor for 4-second epochs. An alternative ('unsmoothed') measure of amplitude (MMDiff) was obtained from the difference between successive maxima and minima generated by a spike detection algorithm in Matlab. A measure of pulse transit time, PTT (between the peaks of the ECG pulses and the next occurring BVP pulses) was also derived from ECG and BVP data processed simultaneously in Matlab [**Fig 3**].

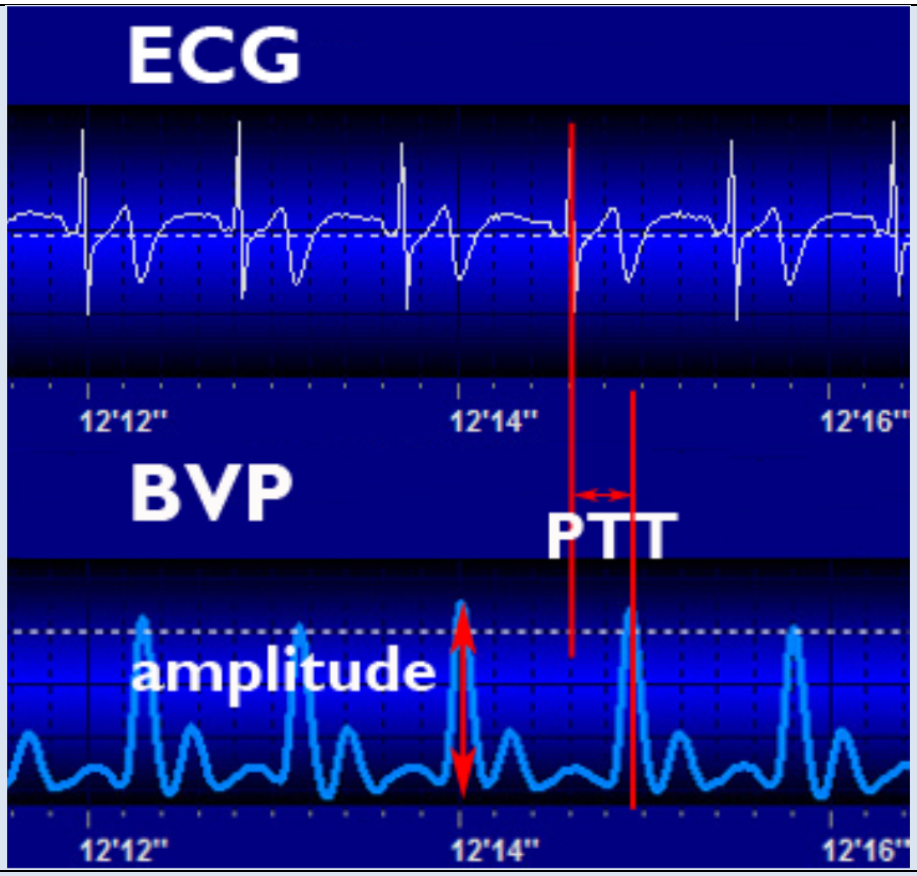


Fig 3. Electrocardiograph (ECG) and Blood Volume Pulse (BVP) traces, showing blood flow amplitude MMDiff and the Pulse Transit Time (PTT) between the ECG 'R' and BVP systolic peaks. Note that in this participant the higher, systolic BVP peak is followed by a notch (the 'dicrotic notch') marking the aortic valve closure, and then a lower peak, the dicrotic wave, marking the diastole.

Analysis: Methods of assessing changes/differences

Changes/differences between the values of a measure in two different 5-minute 'slots' within the same session can be assessed in different ways. For example:

- Direct comparison of the means or medians of the measure for the two different slots [Measure_DIRECT]
- Comparison of the ratios of the two slot means or medians to those of the same measure at baseline (alternatively, instead of ratios, differences from baseline could be compared) [Measure_BASE]
- Comparing means or medians of the ratios (or differences) of the measure in the slots immediately preceding and following the two slots in question [Measure_PRE-POST].

Several of these methods were explored, for both the raw and processed finger temperature and blood flow data (see below).

Analysis: Data pre-processing

Data may be analysed statistically in different ways, depending on whether it is normally distributed or not. Therefore data – split by the levels of the factors to be investigated – was first explored for normality of distribution. This was assessed as acceptable if the result from Shapiro-Wilk test was not significant ($p>0.050$) and the absolute values of skewness and kurtosis divided by their standard error were both <1.96 [Abu-Bader 2010]. Normal Q-Q probability plots were not used as they are not always easy to interpret.

The sort of time series data recorded in this study invariably contains artefacts, whether due to movement or electrical interference. Therefore considerable time was spent exploring artefact correction using graphing methods (Excel) and multiple imputation (SPSS), and assessing to what extent this impacted on the measures investigated (e.g. MMDiff). Some artefact correction was also carried out in Matlab using an accepted algorithm by the researcher experienced in the procedures involved (TS).

In Matlab, the ECG data was processed using Butterworth second order filters (0.01 Hz high-pass and 0.2 Hz low-pass) to remove low frequency drift and high frequency noise or stimulator interference (as in the graph of ECG for participant AI, in **Fig 4** below). Data was then normalised, and Eli Billauer's peak detection method used (peakdet v 3.4.05) [Billauer 2012]. BVP data was similarly processed (using only the low-pass filter), and the PTT extracted from a plot of both the ECG and BVP peaks using a reiterative matrix-based procedure. Care was taken to differentiate the PTT between the ECG R peak and BVP systolic peak from that between the R and diastolic peaks.

Temperature, BVP amplitude and PTT data were examined for outliers in Excel, and analysis conducted on both the 'raw' and the 'processed' data (with outliers/obvious artefacts removed according to simple standard protocol developed for this study). Both the means and medians of these measures and of the raw MMDiff were explored for subsequent analysis. For PTT, modes were also explored.

Results

Technical issues

Four major technical issues were encountered, the first two expected, the other two not.

The presence of artefacts, and their correction

Firstly, as mentioned above, the TEA stimulation introduced noise into the ECG and BVP records during the stimulation slots, with the higher stimulation frequencies leading to denser interference, as in the examples in **Figs 4-6** below. In the first of these (AI, **Fig 4**), stimulation in slot 2 was at 80 Hz, in slot 5 at 2.5 Hz, and in slot 8 at 10 Hz. In the second example (AM, **Fig 5**), stimulation in slot 2 was at 80 Hz, in slot 5 at 10 Hz, and in Slot 8 at 2.5 Hz.

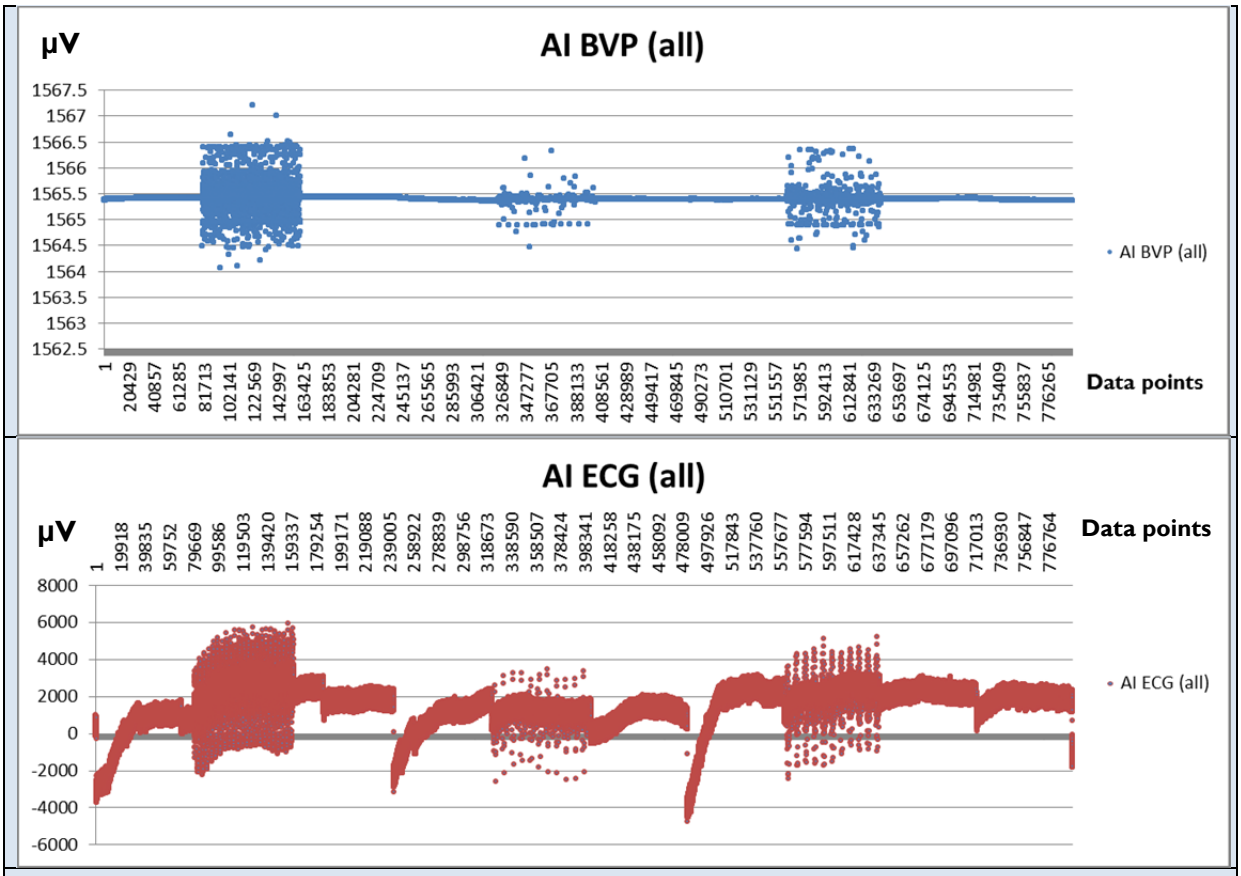


Fig 4. BVP and ECG for participant AI.

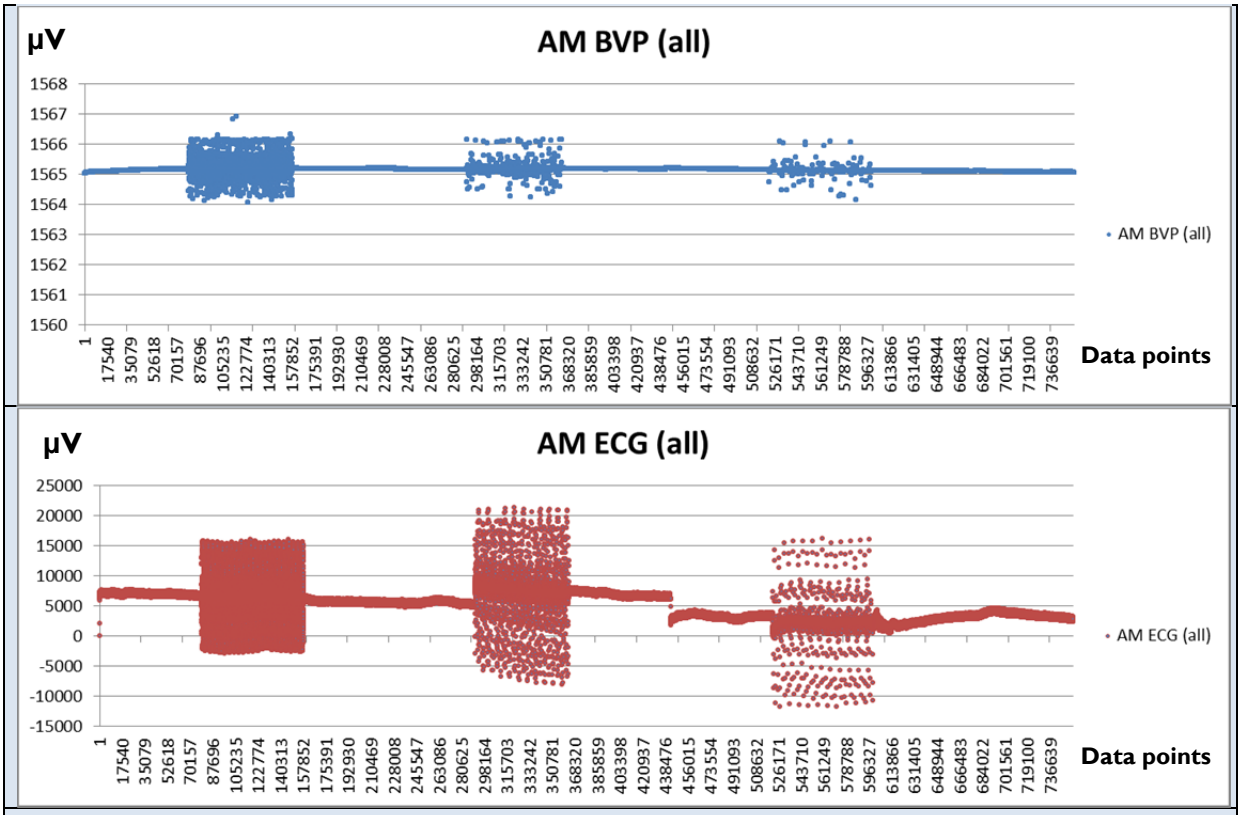


Fig 5. BVP and ECG for participant AM.

However, in some sessions gross interference from stimulation was less evident, as in the BVP record below, for participant AK:

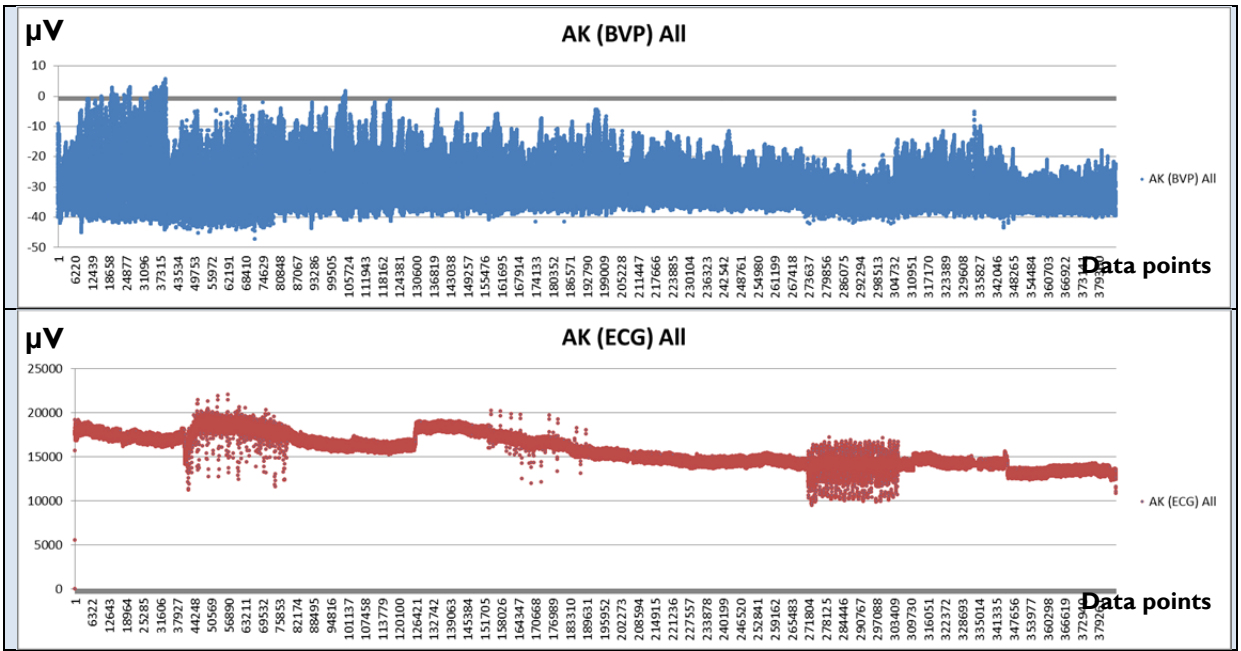


Fig 6A. BVP and ECG for participant AK.

Although the BVP record for participant AK's session might appear at first sight to be more noisy than the other examples, this is in part due to the different scale employed on the y-axis. On zooming in further [**Fig 6B**, below] it does actually show a regular pulse:

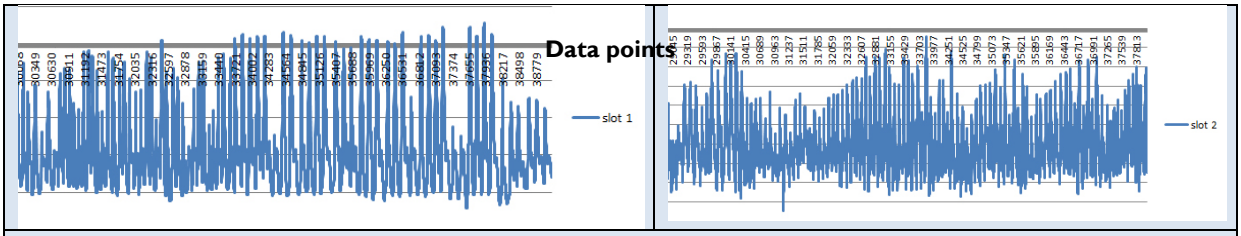


Fig 6B. BVP for participant AK. Close-up view showing regularity hidden in apparent noise.

Artefact correction was carried out in the same way on all the session records, so that amplitude and other signal characteristics would reflect genuine physiological responses rather than the results of data manipulation to remove noise.

Differences between the raw and processed finger temperature data were minimal: for the whole sample (N=12), around 0.03% for the slot means, although as much as 2.7% for the slot medians. Most variance was contributed by two participants (AC, AM).

Corresponding differences for BVP amplitude were around 0.81% for slot means, and 0.58% for slot medians. Most variance was contributed by AE and AI (means), and AC, AN and AQ (medians).

For PTT, the differences were of a different order, with a mean of 11.64% for slot means (10.17% for the medians), although the median differences were lower (7.16% for the slot means, 1.19% for the medians). Data from 14 slots was not recoverable because of the presence of major artefacts, and for the same reason median PTT for a further 12 slots was outside the expected range.

Differences of <3% were considered acceptable in the context of this pilot study.

Secondly, ECG was recorded at a sampling rate of 256 Hz (samples per second), and BVP at 128 Hz. The peak-to-peak difference for PTT was of the order of 80 ms. However, with a sampling rate of 128 Hz, the best accuracy possible was only around 8 ms (1000/128 = 7.8125 ms), and for 256 Hz, around 4 ms (3.9063 ms). Thus, when combining the ECG and BVP recordings, the possible error could be around 12 ms, or 15%.

<u>Other issues</u>

Thirdly, the temperature sensor stopped functioning partway through the series of pre-booked sessions, so that finger temperature readings were not available for 5 participants (29% of the sample to date).

Fourthly, out of a total of 170 5-minute BioTrace+ recording slots, 62 (around 36%) turned out to be shorter, although timed carefully with reference to an external timer (on a separate laptop used to record EEG data). The cause of this data corruption could not be determined, nor was it clear if there was any pattern to the way data had dropped out, or even if the data had been compressed in some way. This fault was intermittent, did not occur during all sessions [**Fig 7A** below], and not necessarily for all slots in a session when it did occur (if anything, more towards the middle or end rather than at

the start of sessions [**Fig 7B** below]. It is suspected that a memory buffer problem in the laptop was responsible, although the wireless Bluetooth link between the NeXus-10 amplifier and the laptop running BioTrace+ may have been faulty in some way. This meant that the number of 'good' (5-minute) slots of the same type that could be compared (stimulation, '5-minute post-stimulation', or the following '10-minute post-stimulation' slots) was limited. In the end, only five sessions consisted solely of 'good' slots. (Unfortunately, not knowing the pattern of data drop-out, it was not possible to overcome the issue by using shorter epochs from the 5-minute slots.)

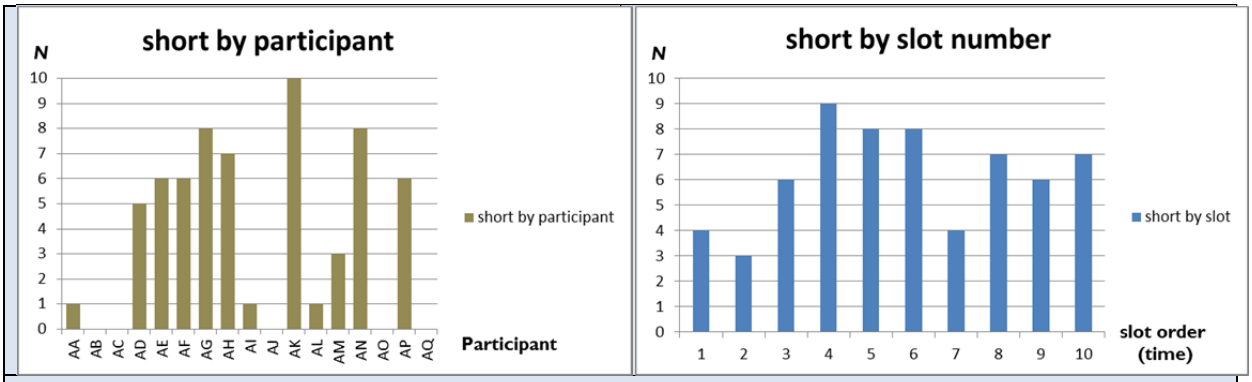


Fig 7. A. Numbers of short slots for each participant; **B.** Numbers of short slots for each of the 10 slots during all sessions.

The 'short slot' technical issue was considered unlikely to impact on finger temperature, a non-temporal measure. However, when the median values of (processed) finger temperature were compared for slots of the correct duration (N=108) and those slots found to be short (N=62), there was a 2.95% difference in the means for the two subgroups [**Fig 8A** below]. Nevertheless, using the nonparametric Mann-Whitney test for independent samples, this difference was not significant.

The corresponding differences for the DIRECT BVP amplitude (median, raw and processed) were less than 1% (Mann-Whitney test n.s.), whereas those for the PRE-POST BVP amplitude were greater than 5%, and for the BASE BVP amplitude as much as 23.96% (p=0.001), definitely unacceptable. For MMDiff, difference between the means for short and correct duration slots was 8.30%, and between the medians 5.22% [**Fig 8B** below] (a non-significant decrease in both cases). For PTT, whether averages, medians or modes were considered, the mean differences were very large (as much as 29.11% for the unprocessed median data), the median differences being somewhat lower (and only 2.92% for the *processed* median data). The Mann-Whitney test showed differences between PTT values in the correct duration and short slots to be highly significant ($p<0.001$), whichever measure of central tendency was used. The large differences for PTT (like those for MMDiff, or between the raw and processed PTT mentioned above) may result in part from its dependence on two sets of data points (ECG and BVP), rather than just a single set, as with BVP amplitude or finger temperature.

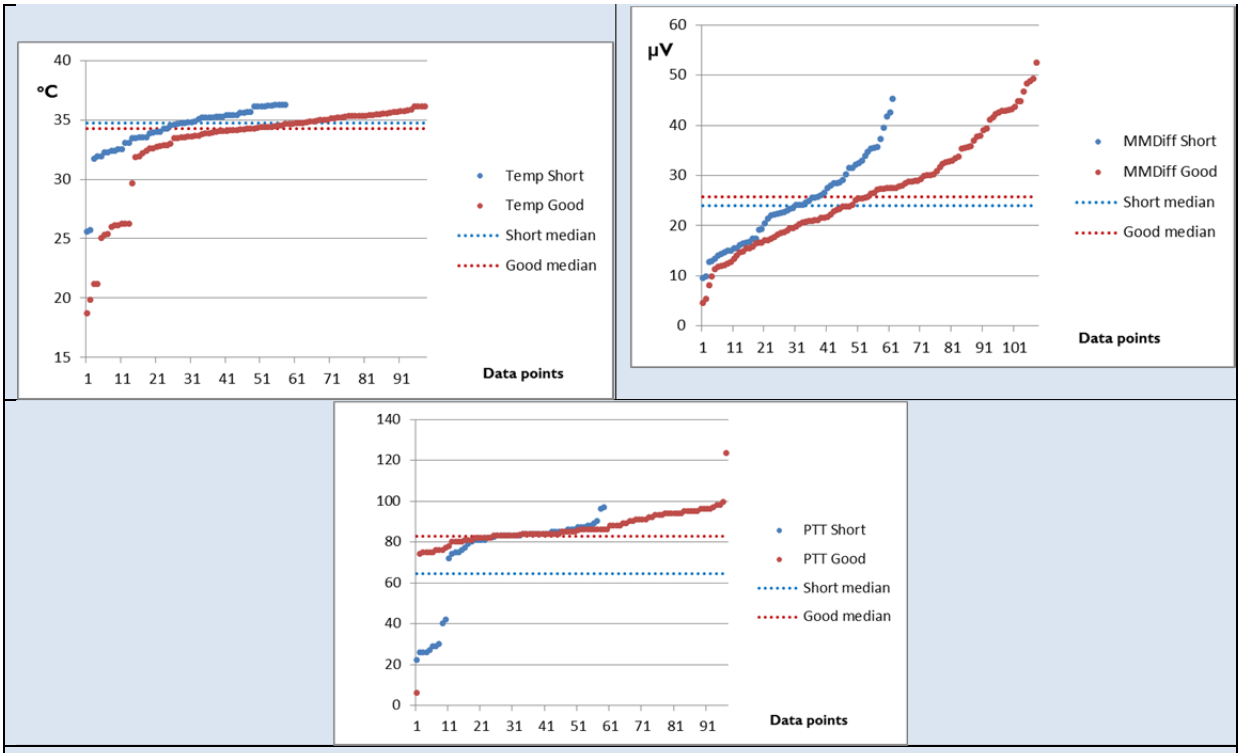


Fig 8. Median slot values over all sessions **A.** Finger temperature. **B.** MMDiff. **C.** PTT. Dotted lines indicate medians for all participants and all slots taken together. Data has been re-ordered from lowest to highest values; that for the short slots is shown in blue, and for the slots of correct duration ('Good') in red.

As can be inferred from the graphs in **Fig 8**, ranges and standard deviations of both temperature and blood flow were greater for the longer slots, but for PTT although the range was greater, SD was smaller.

Fig 9 demonstrates that when comparing session data for the whole sample (N=170) with that for the slots of correct duration considered alone, it appears that removing the short slots affects median finger temperature (median, processed) more during the earlier session slots [**Fig 9A**]. Removing short slots also affects median blood flow (DIRECT BVP amplitude and MMDiff, median processed), but not so much early in sessions [**Fig 9B**]. Removal of short slots affects median PTT in different directions early and late in sessions [**Fig 9C**]. For ease of comparison, all data is normalised to 100 at start of session (except for BVP amp, normalised to 80); note that percentage changes are smaller for finger temperature than for the other measures.

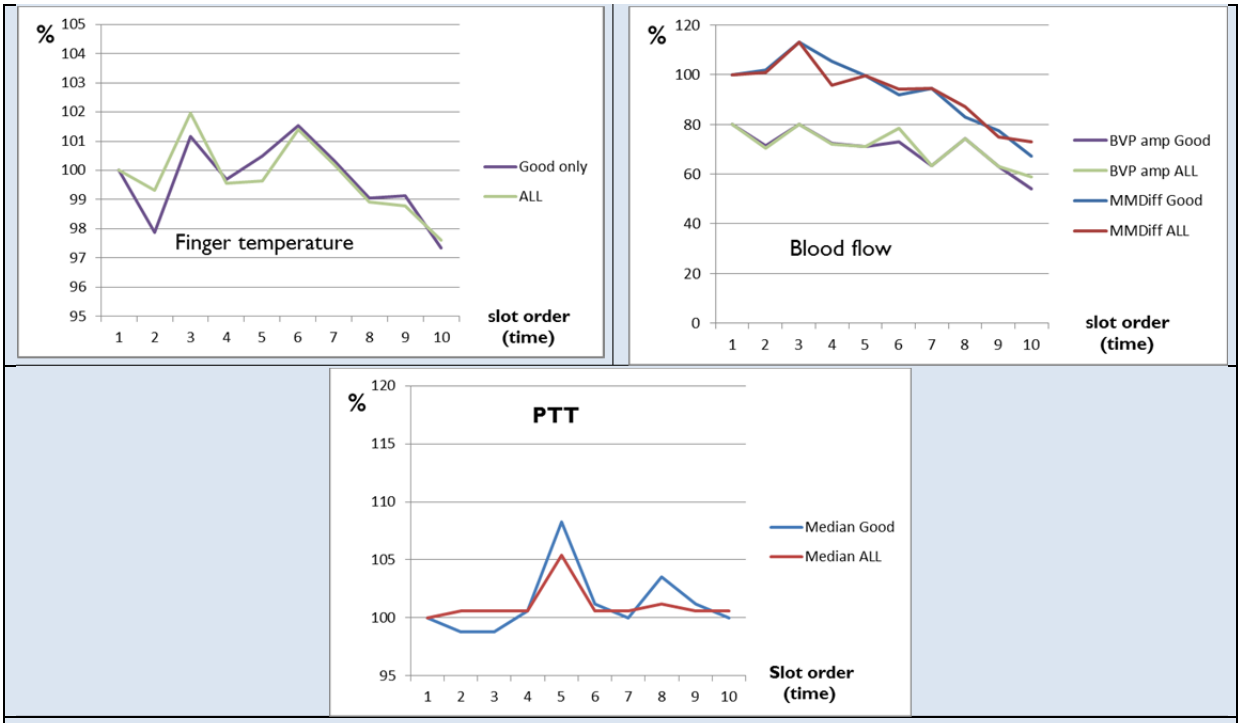


Fig 9. Group medians of changes over session time. For ease of comparison, all data normalised to 100 at start of session (except for BVP amp, normalised to 80). **A.** Finger temperature (median, processed). **B.** Blood flow amplitude, both smoothed (BVP amp) and unsmoothed (MMDiff). **C.** PTT. Differences are visible between the data for 'Good' slots (those of the correct duration) and for 'ALL' slots (Good + short).

Slot duration appears to have less of an impact on DIRECT BVP amplitude data than on MMDiff. Thus further analysis was predominantly carried out on DIRECT finger temperature and BVP amplitudes. For PTT, analysis was only conducted on slots of the correct duration.

Surrogate data

Four artificially shortened slots of the same length were generated from correct data (slot AB4) using different methods, in an attempt to assess how much these would affect data descriptives (see **Appendix IV** for methods used).

For finger temperature, most descriptives were minimally affected, and less than the 2.95% difference between those for the correct and 'real' short slots reported above.

However, the mean MaxMinDiff was around 13% less than that for the original data for all 4 methods used, a relative difference greater than that between correct and 'real' short slots reported above. The normalised standard deviation (nSD, or the ratio of SD to mean) was as much as 13% less for the surrogate than the real data for the first two methods used, but very slightly (1.05%) more for the last method.

Surrogate data was not created for PTT.

Further discussion of surrogate data can be found below, under the heading 'Surrogate data for PRV measures'.

Normality of distribution

Temperature

Finger temperature was not normally distributed, either for the whole dataset or when this was split for various experimental factors (including stimulation frequency).

Blood flow

Of the various direct BVP amplitude measures (mean and median, raw and processed, DIRECT, BASE and PRE-POST), only the DIRECT processed mean and median values were normally distributed for the complete dataset. Neither mean nor median MMDiff were normally distributed (although, if the Kolmogorov-Smirnov test –appropriate for larger samples than available to us – was substituted for the Shapiro-Wilks test, they could appear to be). Thus parametric statistical methods could only be used for some, but not all comparisons – both for the data as a whole, and when split into stimulation, 5-minute post-stimulation and 10-minute post-stimulation slot data. These are shown in **Table 1** below for the whole dataset and for the stimulation and 5-minute post-stimulation slots. (Normality was not assessed for the 10-minute post-stimulation slots. A description of the factors is provided on p 24 under the heading 'Factors impacting temperature and blood flow', and a Key to factor levels in **Table 2**.

Table 1. DIRECT BVP amplitude measures: normality of distribution. If normality was found for unfactored data, this is indicated by 'y'; when the data was split by factor, the factor levels for which normality was found are indicated by numbers, or by 'all' if normality of distribution was found for all levels (no number is given if only one factor was normally distributed). Results when only correct duration slots were checked are shown in square brackets.

Slot	Factor	Flow mean raw	Flow mean processed	Flow median raw	Flow median processed
ALL data	(No factor)				
	Gender				
	Age	1 [all]	[all]	[all]	
	Handedness				
	Prior Rx				
	Hz	[2.5/10 Hz]	2.5/10 Hz [2.5/10 Hz]	[2.5/10 Hz]	[2.5/10 Hz]
	Fav	123 [123]	23 [123]	23 [123]	23 [123]
	LRthresh	[13]	[13]	[13]	
	LRtol	12 [12]	12 [12]	12 [12]	[12]
	Twitch	[012]	[012]	[012]	[012]
Stim slots	(No factor)	y [y]	y [y]	y [y]	y [y]
	Gender	[all]	[all]	all [all]	all [all]
	Age	[all]	[all]	[all]	[all]
	Handedness	[all]	[all]	[all]	[all]
	Prior Rx	[all]	[all]	[all]	[all]
	Hz	2.5/80 Hz [all]	2.5/80 Hz [all]	2.5/80 Hz [all]	2.5/80 Hz [all]
	Fav	013 [all]	013 [all]	013 [all]	013 [all]
	LRthresh	13 [all]	13 [all]	13 [all]	13 [all]
	LRtol	13 [all]	13 [all]	13 [all]	13 [all]

	Twitch	03 [03]	03 [03]	03 [03]	03 [03]
5'min Post-stim slots	(No factor)	y [y]	y [y]	y [y]	y [y]
	Gender	[all]	all [all]	[all]	all [all]
	Age	all [all]	all [all]	all [all]	all [all]
	Handedness	all [all]	all [all]	all [all]	all [all]
	Prior Rx	all [all]	all [all]	all [all]	all [all]
	Hz	2.5/80 Hz [all]	all [all]	all [all]	all [all]
	Fav	013 [13]	13 [13]	013 [13]	013 [13]
	LRthresh	13 [all]	all [all]	13 [all]	all [all]
	LRtol	23 [all]	all [all]	23 [all]	all [all]
	Twitch	[03]	03 [03]	[03]	03 [03]

Table 2. Key to factor levels in Table 1. Counts for each factor are shown in parentheses.

Factor	***N***	**Numbered levels (numbers of cases)**
Gender	2	1. Male (6); 2. Female (11)
Age	2	1. Younger (18-32: 9); 2. Older (38-72: 8)
Handedness	2	0. Right-handed (14); 1. Left-handed (3)
Prior Rx	2	1. Of any of acupuncture; EA or TENS (6); 2. Of none of these (11)
Hz	3	2.5, 10 or 80 Hz (17 of each)
Fav	4	0: Not asked/No comment (31); 1. Most liked (11); 2. Between Most and Least liked (3); 3. Least liked (6)
LRthresh[a]	3	1. High (19); 2. Medium (19); 3. Low (13)
LRtol[a]	3	1. High (18); 2. Medium (18); 3. Low (15)
Twitch	4	0. none (17); 1. Left (4); 2. Right (3); 3. Both Left & Right (27)

a. 'High', 'Medium' and 'Low' defined as within each session, not for the group as a whole.

Interestingly, more BVP amplitude data (factored or otherwise) was normally distributed in the post-stimulation slots than during stimulation or for all slots taken together. Furthermore, more data was normally distributed when only the correct duration slots were considered. However, because *no* factored data was normally distributed both during and after stimulation for any factor and for both the short and correct duration slots, to avoid confusion further analysis was conducted predominantly using either Bootstrap or nonparametric methods.

The PTT data was for the most part not normally distributed.

More about blood flow

The smoothed and unsmoothed measures of blood flow amplitude, BVP amplitude and MMDiff, correlate very closely, but are not identical, as shown in **Fig 10.** Note that the intercept for the scatterplot of the medians [**Fig 10A**] is closer to zero than that for the means [**Fig 10B**].

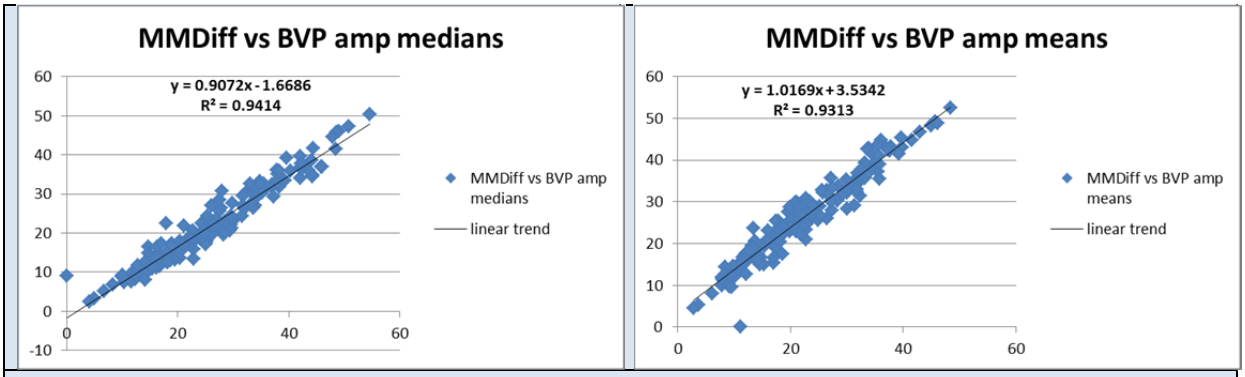


Fig 10. Scatter plots showing correlation between smoothed and unsmoothed measures of blood flow amplitude. **A.** Medians. **B.** Means.

Spearman's *rho* was 0.969 for the whole sample (p<0.001) (1000-sample Bootstrap 95% CI 0.948 to 0.979). For individual participants, *rho* varied between 0.685 (AG; Bootstrap 95% CI 0.000 to 0.987, but with a bias of -0.049) to 1.000 (AL; Bootstrap 95% CI 1.000 to 1.000).

Correlations between temperature, blood flow and PTT

For the whole sample, the association between finger temperature and blood flow (BVP amplitude) was not linear. Thus, even though BVP amplitude was normally distributed, Spearman's ρ (*rho*) was used as a measure of association rather than Pearson's R. Correlations (ρ) of temperature and blood flow were higher between the medians (0.507, 0.507) than between the means (0.436, 0.456) (the first figure shown in each pair is ρ for the raw data, the second for the processed data). Only the medians were considered in further comparisons.

Correlation (ρ) between median temperature and BVP amplitude (both processed) was significant in 7 of the 12 participants for whom finger temperature data was available, with values between 0.673 (p=0.033) and 0.952 (p<0.001). In 10 cases, ρ was positive, and in two negative (although not significant). Thus, in general, finger temperature increased with cutaneous blood flow, as would be expected.

Correlations (ρ) between median temperature (processed) and MMDiff (raw) were similar to those between processed temperature and BVP amplitude, although slightly lower on average (mean ρ 0.465 vs 0.550) [**Fig 11**].

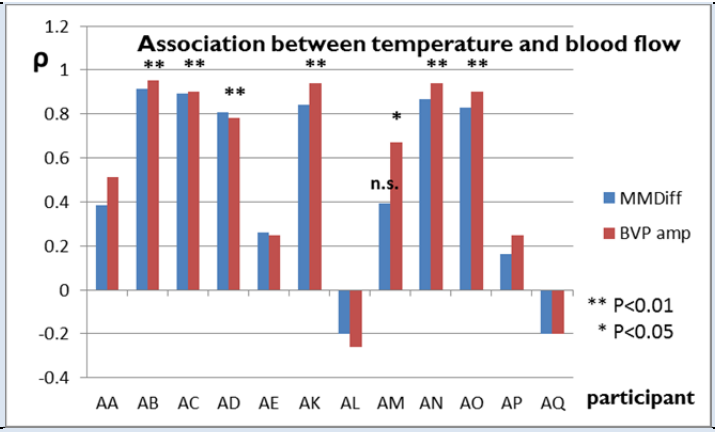


Fig 11. Correlations (ρ) between median temperature (processed) and MMDiff (raw).

For all participants taken together, there was no significant correlation between PTT (median, processed) and MMDiff (raw) or BVP amplitude. The correlation between PTT and finger temperature (median, processed) was, however, significant, and also *negative*, with *rho* = -0.379 ($p<0.001$) for ALL slots (short or of correct duration), and -0.324 ($p=0.006$) for the Good slots (of correct duration only).

There were three participants for whom the correlations between PTT and BVP amplitude was significant (using ALL data), but for only one of these (AC) did the correlation remain significant when only the Good data was analysed (*rho* = 0.833, $p=0.001$). The correlation between PTT and finger temperature was significant only for four participants (ALL data), and for only two when the Good data was considered (AC and AL, with *rho* again >0.8 in both instances).

When correlations are considered for the different time slots rather than by participant, if all 16 possible correlations are considered between the 4 Finger temperature and BVP amplitude measures (mean, median; raw, processed), although the mean of Pearson's R for these is significant in the early slots (slots 1 to 5 for ALL data slots, but only 1 to 4 if only the correct duration, 'Good' slots are considered), it is not significant in the later slots. And Spearman's ρ is significant only for 7 of the comparisons for ALL data in slot 10, not otherwise. The graphs in **Fig 12** below show mean R and ρ for the 10 slots, in time order.

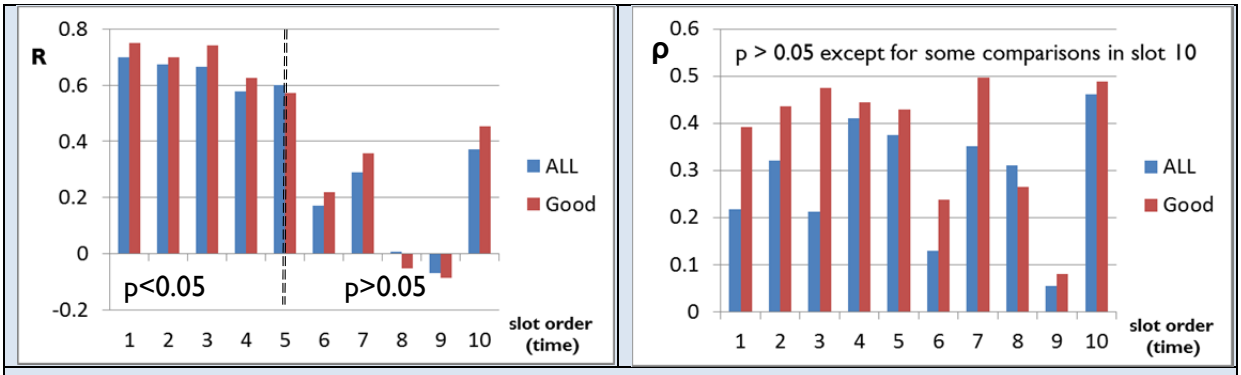


Fig 12. Correlations between finger temperature and BVP amplitude, slot by slot, for all 16 possible comparisons (mean, median; raw, processed), and for ALL and Good data.
A. Values of Pearson's R. **B.** Values of Spearman's ρ.

Thus it would be misleading, in this case, to use Pearson's R, a parametric measure of association, rather than Spearman's ρ, a non-parametric measure more appropriate when the association between finger temperature and blood flow (BVP amplitude) is not linear (above).

Changes over time: Comparisons between temperature, blood flow and PTT

Room, core (oral) and finger temperatures were plotted, together with BVP amplitude and MMDiff for each participant session (10 slots), as in the following example [**Fig 13A**]. PPT for the same session (AL) is shown in **Fig 13B**.

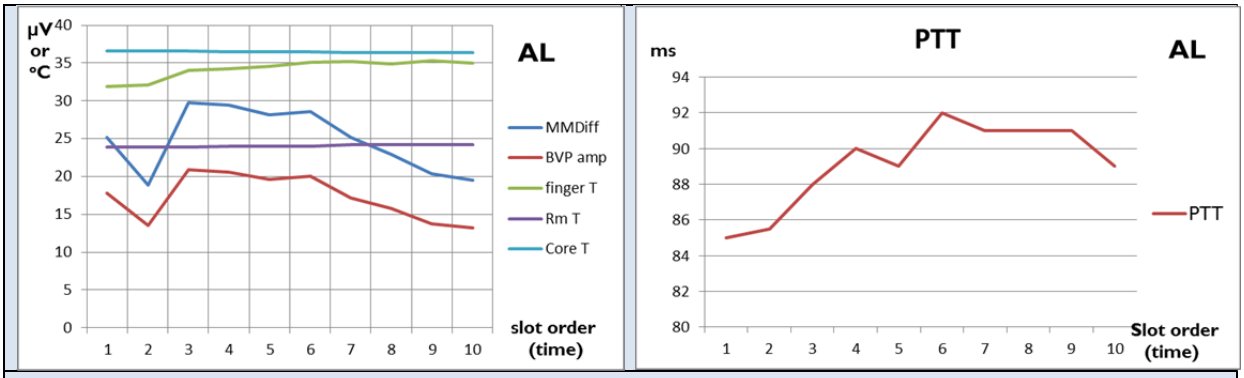


Fig 13. Changes over the session for participant AL. **A.** Temperature and blood flow. **B.** PTT (median, processed).

In most cases, changes over time were very similar for MMDiff and BVP amplitude, although the numerical difference between them varied between participants. In 14 out of 17 sessions, blood flow showed a decrease over the course of the session, and in 8 out of 12 sessions finger temperature did as well – both presumably due to prolonged inactivity; however, as shown in **Fig 11** and **Fig 13A**, finger temperature actually *increased* in two sessions (AL and AQ), during which blood flow decreased. Whereas room temperature remained steady or *in*creased slightly (but did not decrease), core temperature remained steady or *de*creased slightly (sometimes more towards the end of the session). Again, this decrease could almost certainly be attributed to inactivity. Changes in PTT for this participant were quite different [**Fig 13B**], although the overall pattern for both blood flow and PTT was for an increase from slot 2, with a decrease from slot 6 onwards.

Median changes over session time for the complete sample (*N*=17) are shown above in **Fig 9**. Both finger temperature and blood flow increase in slot 3, but by the end of the session have decreased compared to baseline. PTT, in contrast, appears to peak in slot 5, with a subsidiary peak in slot 8, and by the end of the session differs little, if at all, from at baseline. The different patterns of response in these outcome measures warrant further investigation.

Peaks in finger temperature and blood flow appear to occur together (as in **Fig 13**). Out of a total of 15 blood flow peaks in sessions where finger temperature was recorded, this occurred 13 times. However, 4 temperature peaks did NOT coincide with flow peaks. Decreases in finger temperature coincided with flow decreases 10 times (twice in AP and AD, once each in 6 of the other 12 sessions where finger temperature recording was carried out).

A visual impression is that larger peaks in flow (more than around 2 units in **Fig 13**) tended to occur during or immediately following stimulation (which was applied during slots 2, 5 and 8). A count of such peaks where both flow calculation methods (BVP amplitude and MMDiff) gave consistent increases showed that the peaks were distributed in the various slots as in **Fig 14A**, with more following than during stimulation (although participants AF, AH and AQ showed no such peaks, AA showed three, and AI and AJ four; the remainder showed only one or two). This pattern was even clearer for finger temperature. In contrast, there was no clear pattern for PTT [**Fig 14B**].

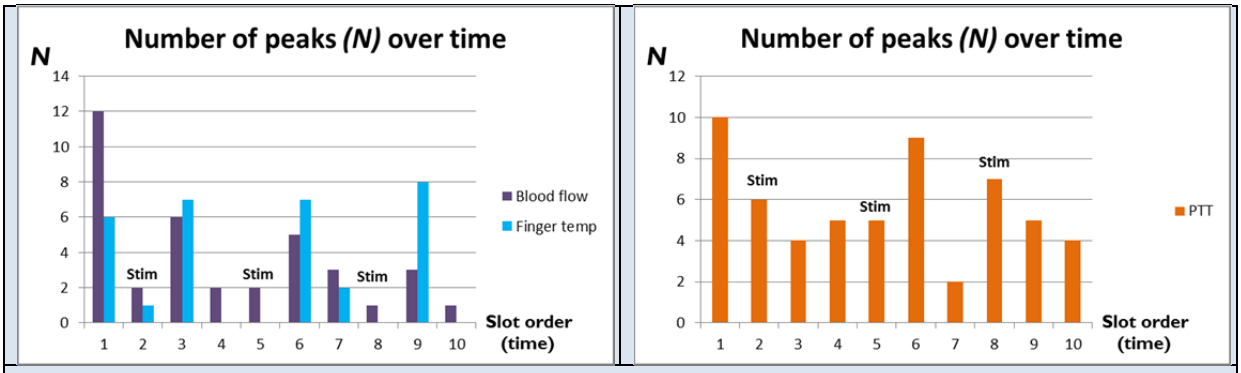
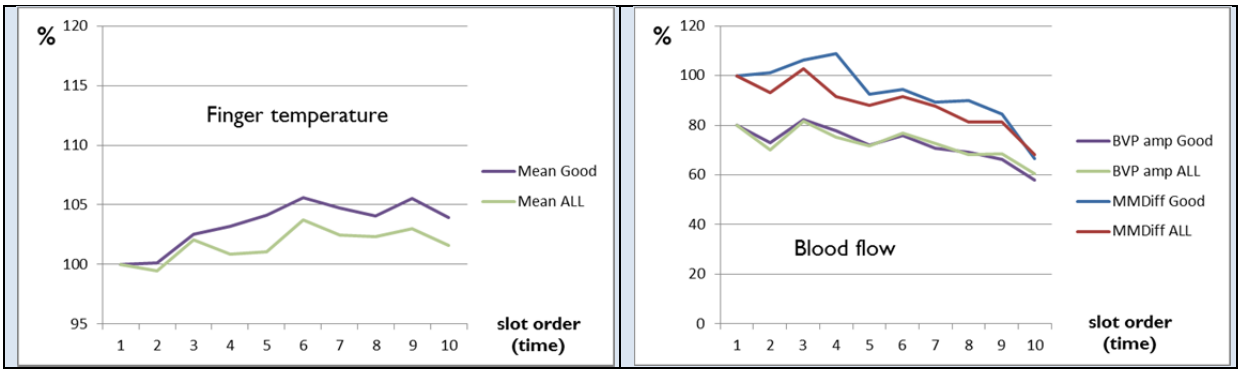


Fig 14. Number of peaks (maxima) per slot for the whole sample to date. **A.** Blood flow (BVP amplitude) and finger temperature. **B.** PTT .

Overall then, as in **Fig 9** above, changes in temperature and blood flow appear to be roughly parallel, but with a different pattern for PTT.

In fact, although the graphs of group *median* values both finger temperature and blood flow (median, processed) both show a decrease over time for all participants taken together[above, **Fig 9**], this is no longer the case for the group *mean* values of finger temperature [**Fig 15A**], although the *mean* values of BVP amplitude and MMDiff still decrease with time [**Fig 15B**]. Furthermore, differences between ALL and Good data for finger temperature now appear *later* rather than early in the session, and are generally more marked for mMDiff. For PTT, the peaks in group median values in slots 5 and 8 [**Fig 9C**] are still visible in the graphs of group means [**Fig 15C**], but otherwise the pattern of change is quite different, with an overall increase over the course of the sessions. (Again, note the small percentage changes in finger temperature compared to those for blood flow and PTT.)



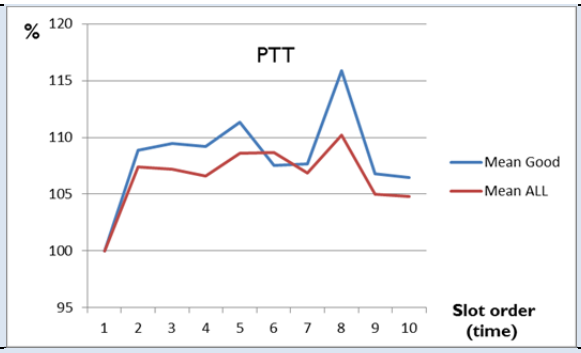


Fig 15. Group means of changes over session time. For ease of comparison, all data normalised to 100 at start of session (except for BVP amp, normalised to 80). **A.** Finger temperature (median, processed). **B.** Blood flow amplitude, both smoothed (BVP amp) and unsmoothed (MMDiff). **C.** PTT.

Great care therefore needs to be taken in selecting from which measures of central tendency (means or medians) conclusions are drawn.

Again, because very few of the outcome measures used – whether of finger temperature, blood flow, PRV or PTT – were normally distributed for all levels of any factor in the different session slots for which these measures were compared, it is sensible to prioritise results based on median rather than mean values, and on nonparametric rather than parametric methods of analysis.

Changes over time: the 'washout' period

An important question is how long the effect of stimulation lasts. Is a 10-minute 'washout' period between 5-minute stimulation slots sufficient, as tentatively assumed in our study protocol, or is there a longer-lasting carry-over effect that will vitiate any findings based on data from subsequent slots?

Consistent changes over 15 minutes

To answer this question, consistency of increases and decreases in outcome measures following a given slot (A, say) over three consecutive slots (B, C, D) were explored. There are eight (2^3) possible patterns of change relative to baseline over three successive slots (from all three increasing to all three decreasing, with intermediate patterns of one increasing and two decreasing, and so forth) [**Table 3**].

Table 3. Possible patterns of change (relative increases/decreases) over three consecutive slots following a given baseline slot.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Pattern	↑ ↑↑	↑↓↓	↑↑↓	↓↓↑	↑↓↑	↓↑↓	↓↑↑	↓↓↓

The final three slots in each session do not have three slots following them, so that, for this group of 17 participants, there are only 7 x 17 (or 119) slots for which patterns of change over the next three slots can be examined. Thus, if all patterns are equally likely to occur, then each would be expected

to appear approximately 119/8 (or 14.875) times, and for any three slots taken together, (3 x 14.875)/8 times, or 5.578 times.

In fact, very surprisingly, there were *no* instances of consistent decreases (↓↓↓) over three successive slots, despite the overall trend for finger temperature and blood flow to decrease over the course of the sessions. Consistent increases over three slots (↑ ↑↑) were as shown in **Table 4**.

Table 4. Consistent increases over three slots following each session slot, in order (maxima in red). Apart from, MMDiff, processed data (median values, ALL slots) was used.

Slot	Finger temp	BVP amplitude	MMDiff (raw)	PTT
1	6	5	5	5
2	8	4	4	6
3	2	1	1	6
4	6	4	6	8
5	6	6	4	5
6	1	2	1	2
7	1	0	0	3
8				
9				
10				
Totals	29	22	21	35

Thus, for each of the four outcome measures, three consecutive increases did occur more frequently than expected, but only for finger temperature and PTT were the differences from 14.875 significant (Binomial, p=0.049 and 0.007, respectively). For PTT, this could be partly the result of the low sampling rate used for ECG and BVP data, and the way PTT was then calculated: the measure may not have been sufficiently sensitive to detect changes (see above, p 11).

If there was a carry-over effect of stimulation, it is likely that the maxima for each measure (in red in **Table 4**) would occur predominantly for the stimulation slots. However, they were evenly distributed between pre-stimulation and stimulation slots. Moreover, a carry-over effect from stimulation would probably result in more occurrences of consistent increases over three consecutive slots following slots 1 (2,3,4)and 4 (5,6,7) than from the triplets of slots following stimulation slots (2, 5) or the 5-minute post-stimulation slots (3, 6), both of which would then include the next stimulation slot (5 and 8). Although there are three more occurrences for MMDiff following slots 1 and 4 than for those following 2 and 5, this is hardly convincing evidence that stimulation has a genuine carry-over effect [**Table 5**].

Table 5. Consistent increases over three slots following pre-stimulation (1,4) and stimulation (2,5) slots.

Slot	Finger temp	BVP amplitude	MMDiff (raw)	PTT
1 and 4	12	9	11	11
2 and 5	14	10	8	11
3 and 6	3	3	2	8

More concerning are the marked differences in occurrence rate of consistent increases following pre-stimulation (1, 4) and post-stimulation (3, 6) slots. These are significant for finger temperature (Binomial test, p=0.035) and MMDiff (p=0.022). This does suggest a **carry-over effect that lasts for 15 minutes**.

Curiously, there appear to be fewer occurrences of consistent increases over three consecutive slots following slot 7 than over three consecutive slots following slots 1 or 4. This may reflect an overall change over the duration of the session rather than the effect of the interventions. As the slots following 3 and 6 are later in the session than those following 1 and 4, the apparent carry-over may be in part an artefact of the effect of time and habituation.

Table 6 shows the results of grouping occurrences of consistent increases over three consecutive slots by stimulation frequency, with 1 and 4 as pre-stimulation slots re-labelled with the frequency applied in the following stimulation slots.

Table 6. Results of grouping occurrences of consistent increases over three consecutive slots by stimulation frequency.

2.5 Hz	**Finger temp**	**BVP amplitude**	**MMDiff (raw)**	**PTT**
1 and 4	4	3	3	2
2 and 5	5	3	3	5
3 and 6	1	1	1	4
4 and 7	3	1	2	2
10 Hz				
1 and 4	3	2	4	5
2 and 5	3	4	3	1
3 and 6	1	1	0	0
4 and 7	1	1	3	4
80 Hz				
1 and 4	6	5	4	6
2 and 5	6	3	1	5
3 and 6	1	1	1	4
4 and 7	3	2	1	5

Again there do not appear to be more occurrences of consistent increases over three consecutive slots following pre-stimulation than stimulation slots, except for PPT for stimulation at 10 Hz, and perhaps BVP amplitude for stimulation at 80 Hz. Nor do there appear to be more consistent increases following pre-stimulation than post-stimulation slots. However, the difference in occurrence rates of consistent increases following pre- and post-stimulation slots is concerning, particularly at 80 Hz (for finger temperature, BVP amplitude and MMDiff) and 10 Hz (for PTT).

Consistent changes over 10 minutes – increases

There are four (2^2) combinations of increases and decreases over two consecutive slots, with 8 x 17 (or 136) slots for which these patterns of change can be examined. Thus, if all patterns are equally likely to occur, then each would be expected to appear approximately 136/4 (or 34) times, and for any three slots taken together, (3 x 34)/8 times, or 12.75 times. Consistent increases over two slots (↑ ↑) were as shown in **Table 7**.

Table 7. Consistent increases over two slots following each session slot, in order (maxima in red). Apart from, MMDiff, processed data (median values, ALL slots) was used.

Slot	Finger temp	BVP amplitude	MMDiff (raw)	PTT
1	5	3	4	3
2	10	9	7	7
3	4	1	1	5
4	6	3	5	5
5	6	6	6	7
6	1	1	3	1
7	2	2	2	2
8	3	1	1	5
9				
10				
Totals	37	26	29	35

Thus, for each of the four outcome measures, two consecutive increases did occur marginally more frequently than expected, but only for finger temperature and PTT (Binomial, n.s.). Consistent increases occurred most commonly for all four measures for the two slots following the first stimulation, but also quite frequently occurred for the two slots following the second stimulation. **This does suggest a carry-over effect lasting for 10 minutes following the novel experience of a first stimulation**, decreasing after the second stimulation slot, and then again, considerably for finger temperature and blood flow, after the final stimulation slot Again, there appear to be fewer occurrences of consistent increases over two consecutive slots following slots 7 and 8 than over two consecutive slots following slots 1 and 2 or 4 and 5. This may reflect an overall change (habituation) over the duration of the session rather than the effect of the interventions themselves.

Table 8A, demonstrates that there is a more consistent 10-minute response following stimulation than following pre-stimulation slots. This difference is not significant (using the Binomial test), nor is that between the actual and expected number (12.75) of consistent decreases.

Table 8A. Consistent increases over two slots following pre-stimulation (1,4,7) and stimulation (2,5,8) slots.

Slot	Finger temp	BVP amplitude	MMDiff (raw)	PTT
1, 4, 7	13	8	11	10
2, 5, 8	19	16	14	19

Table 8B demonstrates that there are significantly more occurrences of a consistent 10-minute response following stimulation than post-stimulation slots (using the Binomial test, for finger temperature p=0.027, for BVP amplitude p=0.002, and for MMDiff p=0.049). However, none of these numbers are significantly different from the expected number of consistent increases (8.5 for 2 slots considered together).

Table 8B. Consistent increases over two slots following pre-stimulation (1,4), stimulation (2,5) and post-stimulation (3,6) slots.

Slot	Finger temp	BVP amplitude	MMDiff (raw)	PTT
1 and 4	11	6	9	8
2 and 5	16	15	13	14
3 and 6	5	2	4	6

However, as noted above, these differences may reflect an overall change over the duration of the session rather than the effect of the interventions, possibly due to habituation (see too **Fig 16**).

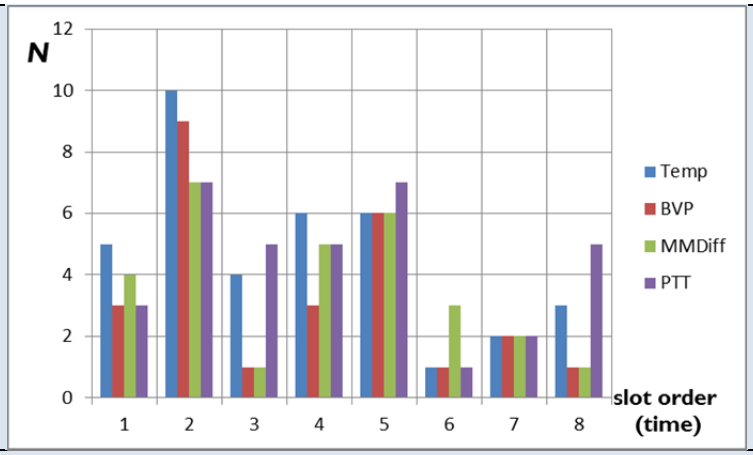


Fig 16. Occurrence of consistent increases over two consecutive slots following each slot (1 to 8).

Table 9. Results of grouping occurrences of consistent increases over two consecutive slots following two slots of the same type, by stimulation frequency.

2.5 Hz	Finger temp	BVP amplitude	MMDiff (raw)	PTT
1 and 4	4	2	2	3
2 and 5	2	0	0	1
3 and 6	2	0	0	0
4 and 7	4	1	2	0
10 Hz				
1 and 4	3	1	3	0
2 and 5	2	3	2	1
3 and 6	2	1	1	0
4 and 7	2	3	3	2
80 Hz				
1 and 4	4	2	0	0
2 and 5	3	4	2	2
3 and 6	1	0	0	2
4 and 7	2	0	0	0

Here numbers are too low to show particular trends, although there do appear to be more occurrences of consistent increases over two consecutive slots following pre-stimulation than stimulation slots at 2.5 Hz (except for PPT). The reverse is true at 80 Hz (except for finger temperature), following which there are more consistent increases than after the 5-minute and 10-

minute post-stimulation slots. If the results for all three slots of the same type are combined, then the differences between occurrences of consistent increases over two consecutive slots following pre-stimulation than stimulation slots at 2.5 Hz increases slightly for all except PTT [**Table 10**]. This suggests that there is a fairly immediate effect of stimulation, but is otherwise difficult to interpret.

Table 10. Results of grouping occurrences of consistent increases over two consecutive slots following three slots of the same type, by stimulation frequency (1, 4 and 7, as pre-stimulation slots, have been re-labelled with the frequency applied in the following slots).

2.5 Hz	**Finger temp**	**BVP amplitude**	**MMDiff (raw)**	**PTT**
1, 4 and 7	5	4	4	3
2, 5 and 8	2	0	0	2
10 Hz				
1, 4 and 7	4	1	3	0
2, 5 and 8	2	3	2	1
80 Hz				
1, 4 and 7	4	2	0	0
2, 5 and 8	3	4	2	2

Consistent changes over 10 minutes – decreases

Table 11 shows consistent decreases (↓↓) over two slots following each session slot, in order.

Table 11. Consistent decreases over two slots following each session slot, in order (maxima in red). Apart from, MMDiff, processed data (median values, ALL slots) was used.

Slot	**Finger temp**	**BVP amplitude**	**MMDiff (raw)**	**PTT**
1	7	8	8	6
2	6	4	7	8
3	13	9	12	7
4	8	9	8	5
5	7	6	6	7
6	14	12	14	12
7	12	11	10	8
8	9	10	10	9
9				
10				
Totals	76	69	75	62

Consistent decreases over two consecutive slots were expected to occur 34 times for each measure (above, p 23), so their actual incidence is significantly different from expected (Binomial test, $p<0.001$, $p=0.001$, $p<0.001$, $p=0.006$, respectively).

Grouping the results together, there appear to be slightly more consistent decreases following pre-stimulation than after stimulation slots, but considerably more decreases following post-stimulation than stimulation slots. Neither of these findings suggests that stimulation in itself contributes to consistent 10-minute decreases in the measures.

Table 12 shows the consistent decreases over two slots following pre-stimulation (1,4), stimulation (2,5) and post-stimulation (3,6) slots. For finger temperature, the difference in occurrence of

consistent decreases following post-stimulation than stimulation slots is significant (Binomial test, p=0.038).

Table 12. Consistent decreases over two slots following pre-stimulation (1,4), stimulation (2,5) and post-stimulation (3,6) slots.

Slot	Finger temp	BVP amplitude	MMDiff (raw)	PTT
1 and 4	15	17	16	11
2 and 5	13	10	13	15
3 and 6	27	21	26	19
4 and 7	20	20	18	13

Table 13 shows the results of grouping occurrences of consistent decreases over two consecutive slots following three slots of the same type, by stimulation frequency (1, 4 and 7, as pre-stimulation slots, have been re-labelled with the frequency applied in the following slots). Here most differences between occurrences of consistent decreases following pre-stimulation and stimulation slots are for 10 Hz, with more decreases following the pre-stimulation slots (except for PTT). As before, this suggests that there is possibly a fairly immediate effect of stimulation at 10 Hz.

Table 13. Results of grouping occurrences of consistent decreases over two consecutive slots following three slots of the same type, by stimulation frequency.

2.5 Hz	Finger temp	BVP amplitude	MMDiff (raw)	PTT
1, 4 and 7	8	8	8	9
2, 5 and 8	7	6	6	5
10 Hz				
1, 4 and 7	10	12	10	5
2, 5 and 8	7	7	9	12
80 Hz				
1, 4 and 7	9	8	8	5
2, 5 and 8	8	7	8	7

Differences from baseline in pre-stimulation slots

A third way of testing whether the 10-minute washout period is sufficient is to check if there are significant differences in values from baseline (slot 1) during pre-stimulation slots 4 and 7 [Alla Machanova, personal communication, 17 March 2015]. Both Friedman and Wilcoxon tests indicated no significant differences for finger temperature between these three slots (considered together, or pair-by-pair), for either ALL (*N*=17) or Good (*N*=6) data. For BVP amplitude (ALL slots), the Friedman test indicated a significant difference (p=0.035), with the Wilcoxon test showing significant differences between values in slots 7 and 4, and slots 7 and 1 (p=0.028 for both comparisons). For BVP amplitude (Good slots), neither the Friedman test nor the Wilcoxon test showed significant differences (although p=0.050 for slot 4 vs slot 7).

Thus, for the Good data, this method of testing for carry-over suggests that values of finger temperature and BVP amplitude blood flow (median, processed) are comparable to baseline by 10 minutes post-stimulation.

Conclusions on washout and carry-over effects

In conclusion, there may be 10-minute or even 15-minute carry-over effects following 5 minutes of stimulation, particularly if this stimulation is novel. However, these effects may reduce rapidly as the stimulation (or situation) becomes more familiar, and the third method used to test for carry-over effects does suggest that – for the six cases of Good data, at least – results for finger temperature and BVP amplitude are comparable to baseline by 10 minutes post-stimulation.

However, especially given the variability of BVP amplitude at baseline (slot 1), this conclusion can only be tenuous. Further investigation is necessary to assess if particular individuals show longer-lasting carry-over effects, and also to gather further data that is not corrupted. Computation to assess carry-over effects on PRV measures (below) has not yet been undertaken, and could help to clarify how important carry-over is..

It is unclear whether the effects of repeating stimulation for 5 minutes every 15 minutes are cumulative, or whether each successive 5 minutes of stimulation 'resets' outcome measures to baseline in some way. Hopefully, because the order at which the three frequencies was applied varied in each session, this will help to overcome any interaction effects due to carry-over or simply the passing of time.

Changes over time: subjective responses

As described above, two subjective measures of mood and feeling were used during experimental sessions, a short form of the Profile of Moods State (POMS-SF) and a Composite NRS for current relaxation and mood. Participants completed the POMS-SF questionnaire twice, before sensors and stimulation electrodes were attached, and again after they were removed. The NRS was completed at these times, and also at around five minutes after each period of stimulation, so five times in all.

Fig 17 shows changes in POMS-SF [**Fig 17A**] and mood NRS [**Fig 17B**] over the course of sessions (N=17).

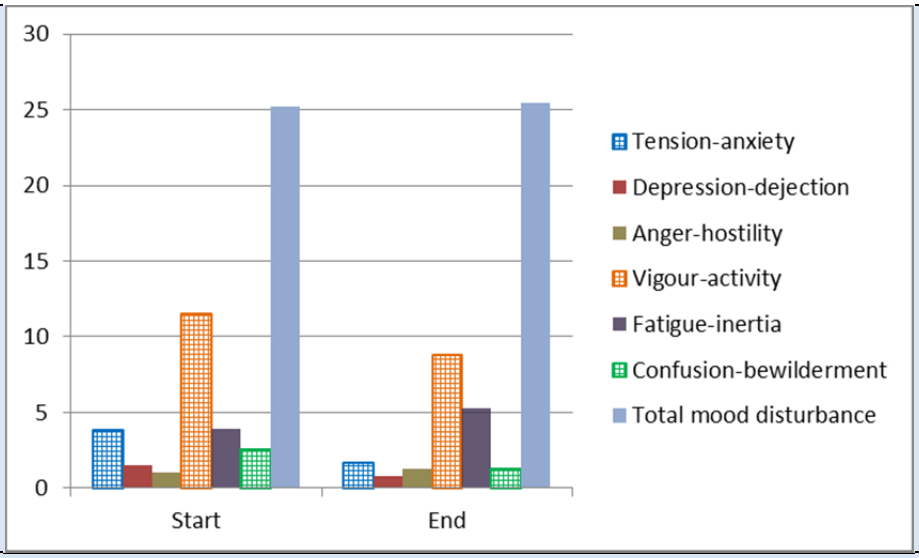


Fig 17A. Changes in POMS-SF scores for current mood over the course of sessions (*N*=17). The POMS-SF was administered twice, at 'Start' and 'End'.

Chequerboard patterning indicates a significant difference in pre and post scores ($p<0.05$ for the Wilcoxon signed-ranks test)

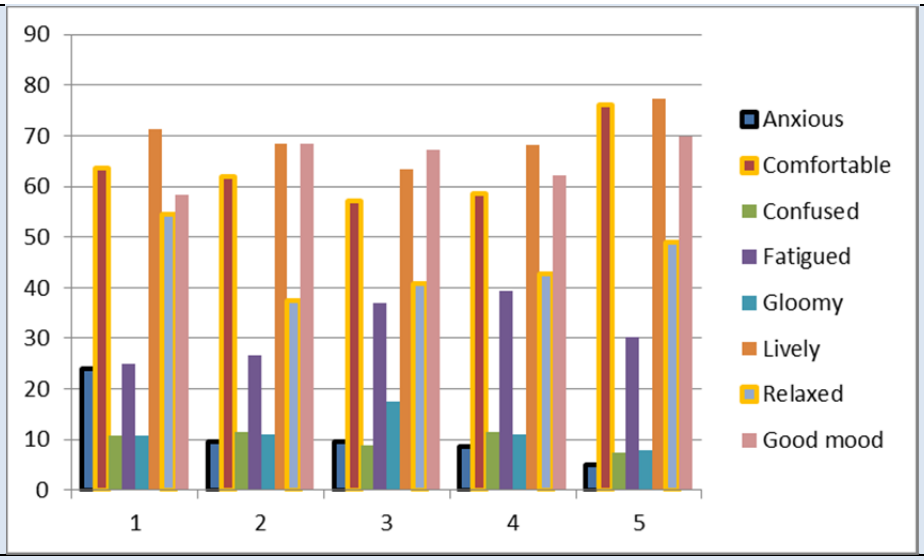


Fig 17B. Changes in NRS scores for current mood over the course of sessions (*N*=17). The NRS was administered five times: at 'Start' and 'End', and before each 10-minute Post-stimulation slot.

Outlined columns indicate a significant difference among the 5 NRS scores (black $p<0.001$, orange $p<0.01$ for the Friedman test).

Note that using both scales there is a significant overall decrease in anxiety over the course of sessions, as well as a decrease in confusion (significant in the POMS-SF but not in the NRS). Changes in feeling 'Relaxed' are significant in the NRS, with participants more relaxed at the start than at the end, although scores recover throughout following a large drop after the first stimulation slot. Feeling 'Comfortable' also changes significantly in the NRS, participants' comfort decreasing towards mid-session and then rebounding once participants are released from encumbering electrodes and having to remain seated. Correspondingly, Fatigue increases – although not significantly – in both scales (decreasing at the end in the NRS), with 'Vigour-activity' also decreasing (significantly); in contrast, although feeling 'Lively' does decrease initially, it then recovers to above baseline. Feeling 'in a good mood' also increases overall in the NRS, but Total mood disturbance (i.e. feeling bad) in the POMS-SF also does, if very slightly.

More significant differences occurred between the second and fifth (4), the fourth and fifth (4), and third and fifth (3) NRS scores, than between other paired NRS scores (0, 1 or 2 comparisons significant from the Wilcoxon signed-ranks test). In fact, the fifth NRS scores were significantly different from the others in 30 out of 32 possible comparisons, whereas the second NRS scores, for example, were significantly different in only 19 out of 32 (the first NRS scores in 24, third in 27 and fourth in 22). This preponderance for the fifth NRS scores is easily explained by the fact that participants were no longer constrained by EEG cap and electrodes, or having to sit still, in contrast to their situation when completing the earlier NRSs.

Appendix III shows the correlations found between the NRS and POMS-SF scale, and between the NRS and Cohen's Perceived Stress scale (PSS-10).

Associations between subjective state and temperature or blood flow

Few of these associations are significant. Those that are are positive, and significant at the 5% rather than the 1% level.

There are 14 significant associations with baseline POMS (and none with the second POMS administered). A higher baseline measure of POMS Tension-anxiety (POMS-T) is associated with greater blood flow (BVP amplitude) in slots 2, 4 and 5 (*rho* between 0.524 and 0.596). Otherwise the POMS associations are all with finger temperature: greater baseline POMS Fatigue-inertia (POMS-F) is associated with higher temperature in slots 4 to 7 (*rho* between 0.577 and 0.608); more baseline POMS Confusion-bewilderment (POMS-C) is associated with higher temperature in slots 6 and 8 (*rho* o.579, 0.586); and Total mood disturbance (POMS-TOT) is positively associated with temperature in slots 5-8 and 10 (*rho* between 0.577 and 0.683).

There are fewer significant associations between NRS items and blood flow (1) or finger temperature (6). The single association with blood flow (BVP amplitude) is between NRS 'Lively' between slots 3 and 4 and blood flow during slot 4 (*rho* 0.501). NRS 'Anxious' at baseline correlates with temperature in the final slot (*rho* 0.636), and between slots 3 and 4 with temperature in slot 3 (*rho* 0.625); NRS 'Confused' also correlates with temperature at this juncture (*rho* 0.609), and between slots 6 and 7 with temperature in slot 6 (as well as in slots 4 and 5: *rho* 0.577 to 0.592).

Apart from the single association with NRS 'Lively', all the other POMS and NRS correlations suggest that skin temperature and blood flow *increase* with negative emotions. This is somewhat counter-intuitive, indicating further investigation is necessary. Associations between subjective state and PTT have not yet been examined.

Correlations between PRV indices and temperature or blood flow

In previous pilot studies, we have investigated the effects of EA and TEA on the following HRV/PRV indices, calculated from the R-R inter-beat intervals (IBIs) derived from BioTrace+ recordings and then analysed in Matlab using a spike detection algorithm [Billauer 2012]:

Time domain (3)

- RR Mean R-R, or beat-to-beat, interval (ms)
- SDNN R-R standard deviation (ms)
- HR Heart rate (bpm)
- RMS SD Root mean square of successive RR differences (ms)

Frequency domain (FFT spectrum using Welch's periodogram) (2)

- HFpwr HF power (mA^2)
- LF/HF LF/HF power ratio,

Nonlinear (3)

- ApEn Approximate entropy
- SampEn Sample entropy
- D2 Correlation dimension.

Regardless of whether slots were of the correct duration or shortened, the correlations shown in **Table 14** appear robust (+ positive, - negative correlation). The association between PTT and PRV measures should also be investigated when time allows (a correlation between SDNN and arterial stiffness has been reported by other authors [Jaiswal et al. 2013]).

Table 14. Correlations between PRV measures and finger temperature/ blood flow.

PRV measure	ALL data	Stim slots	5-min Post	1-min Post
RR	BVP amp (all) (D & B) [-]		BVP amp (all*) (D) [-]	BVP amp (all) (D & B) [-]
SDNN	BVP amp (all) (D) [-]	BVP amp (all) (D) [-]		Finger temp (all) [-]
HR	BVP amp (all) (D & B) [+]			
RMS SD	BVP amp (all) [-]			
HFpwr				
LF/HF				
ApEn	BVP amp (all) (D) [+]	BVP amp (all) (D) [+]		
SampEn				
D_2	BVP amp (all) (D) [-]	BVP amp (all) (D) [-]		

Note: D=DIRECT, B=BASE, P=POST-POST. * Except for raw median BVP amplitude.

The RR and HR correlations suggest that as heart rate decreases (and the RR interval increases accordingly), blood flow also decreases. This makes intuitive sense. SDNN and RMS SD increase with PRV variability, so like RR could be expected to increase as blood flow decreases. However, that the correlations of blood flow with ApEn and D_2 (different ways of assessing variability) are in opposite directions is not so easy to understand.

What about the correlations which were not significant for the remaining three measures: HFpwr, LF/HF and SampEn?

Taken together over the course of all the sessions, mean HFpwr decreases, whereas median HFpwr increases (for both ALL slots and only those of the correct duration). For LF/HF, these changes are reversed [**Fig 18**]. Given that an increase in HFpwr and a decrease in LF/HF are considered to represent increases in parasympathetic activity (relaxation)or reduced sympathetic modulation [Malik 1996], if the mean values are considered, this would suggest that relaxation *decreases* over the course of a session. In contrast, the median values would suggest an *increase* in relaxation.

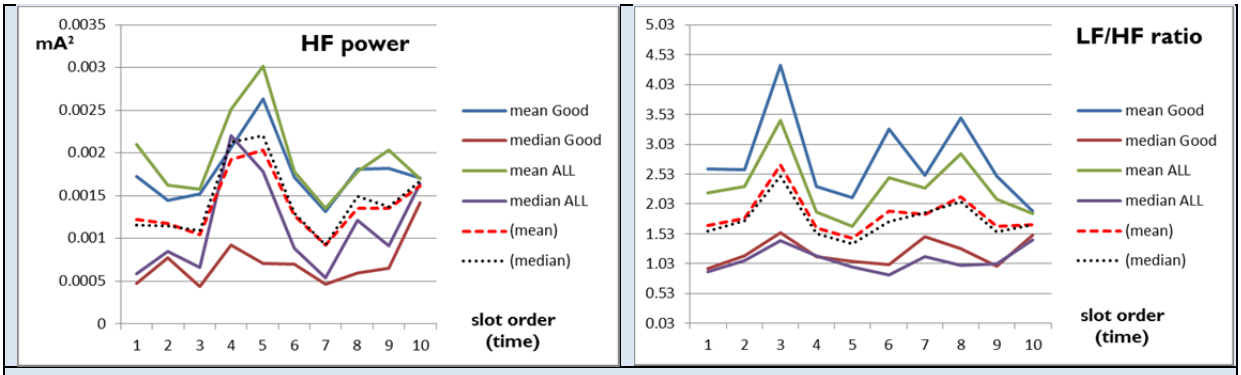


Fig 18. Mean and median values of PRV measures over time, for ALL and Good slots. **A.** HF power. **B.** LF/HF power ratio.

Thus the decreased blood flow over the course of the sessions may, paradoxically, be associated with an increase in parasympathetic tone, or a decrease in sympathetic modulation. This corresponds to the counterintuitive finding (above, p 30) that skin temperature and blood flow may *increase* with negative emotions.

For SampEn, both mean and median increase if only the slots of correct duration are considered, but decrease if data is taken from both the good and short slots [**Fig 19**]. Given that increases in SampEn are again likely to reflect better cardiovascular health (and parasympathetic function) [Steffert & Mayor 2014], this confirms that the short slot data is unlikely to give an accurate indication of cardiovascular function and should only be used with extreme caution.

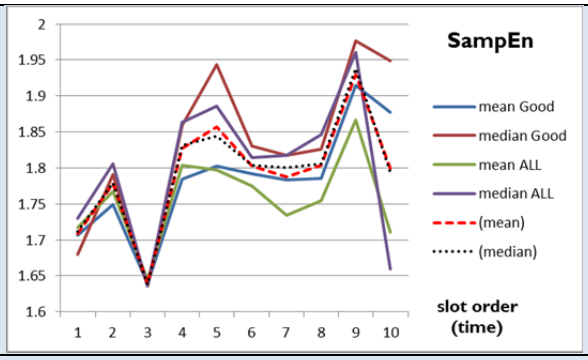


Fig 19. Mean and median values of PRV measures over time, for ALL and Good slots. Sample entropy.

Comparing these graphs with those for finger temperature and blood flow [**Fig 15**], it is evident that the time-dependent PRV measures (time and frequency domain) are more affected by the data drop-out/compression issue than are finger temperature and blood flow.

Table 15 shows the percentage differences between the PRV indices of the short slots and slots of the correct lengths: %Diff = (short – correct)/correct%.

Table 15. Percentage differences between the PRV indices of the short slots and slots of the correct lengths.

%Diff	RR	SDNN	HR	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D2
mean	-30.88	37.43	78.12	46.39	36.57	-53.72	-1.61	-4.09	8.25
median	-26.97	94.00	41.22	110.77	212.18	-11.28	-3.76	-0.59	14.65

The nonlinear PRV measures Approximate entropy (ApEn), Sample entropy (SampEn) and Correlation dimension (D_2) are less affected by the data drop-out/compression issue than the others. However, being more stable may also mean that these particular measures are less useful for detecting differences between, for example, stimulation at different frequencies. They are also less easy to interpret physiologically.

Surrogate data for PRV measures

A. Using the surrogate data described above (p 14), percentage differences between the surrogate and actual data for slot AB4 were as shown in **Table 16**.

Table 16. Percentage differences between the surrogate and actual data for slot AB4.

%Diff	RR	SDNN	HR	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D2
method1	-1.99	4.62	2.18	34.25	333.61	-86.06	-6.29	54.43	10.80
method2	-3.38	32.04	4.48	50.11	633.06	-89.94	7.55	54.95	-85.32
method3	-3.07	27.90	3.92	53.14	623.95	-88.01	4.23	58.94	-95.94
method4	-2.53	31.58	3.52	70.46	594.58	-89.66	6.18	73.52	-97.22

These percentages are rather different to those in **Table 15**, suggesting that the actual data drop-out/compression occurring in the short slots did not follow anything like the patterns adopted in creating the surrogate short slot versions of AB4.

B. For each of the 62 short slots, the ratio of 'compressed' to actual duration was different, varying between 0.168 and 0.977 (median 0.708; mean 0.709, SD 0.210). For each slot, the IBIs were then divided by this ratio, and the new IBIs created used to generate surrogate values for the PRV measures in Kubios HRV. When the medians of these PRV measures were compared with those for the correct duration slots, RR, HR and ApEn were within 20% of the 'true' values that could be expected for the short slots, and SampEn within 30% of the expected value (apart from participant AL). However, attempting to correct the short slot data using this method did not work well for the other measures; nor did it work well for some participants (e.g. AI, AL, AN, AP). Thus uniform data compression does not account for the existence of the short slots.

Factors impacting temperature, blood flow and PTT

It is very difficult to design an experiment in which all possible factors except the one being investigated are kept constant. In this pilot study, in addition to stimulation frequency the following were considered:

- Participant
- Age
- Gender
- Prior experience of acupuncture, EA, TEA or any of these
- Handedness
- Room temperature
- Time (order of slots)
- Stimulation amplitude sensory or tolerance threshold (left, right or the average of both)
- Occurrence of twitch in response to stimulation (left, right, left and right)
- Most or least favourite of the three stimulation slots.

The correlation ratio *eta* (*η*) was computed in SPSS in order to rank the effect of these factors for all data taken together, as in a previous pilot study in this series [Steffert & Mayor 2014]. In the lists below, greatest variation was found for those factors at the top of the list, least for those at the bottom. For finger temperature, which was only minimally processed for artefacts, *eta* (*η*) was the same for both raw and processed data, and for most factors marginally (not significantly) lower for medians than means. For DIRECT and PRE-POST BVP amplitudes, *η* was in general greater for processed than for raw values, and for mean than for median values, whereas this was reversed for BASE.

Table 17 shows the experimental factors ranked by value of correlation ratio *eta* (*η*). For the 12 possible BVP amplitude measures, a 'compound eta count' (CEC) was also used to roughly rank each factor:

$$CEC = \Sigma([N \text{ factor occurrences for which } \eta > 0.n] \times n \qquad (\text{where } 1 \le n \le 9)$$

Table 17. Experimental factors ranked by value of correlation ratio *eta* (*η*). Factors for finger temperature for which *eta* (η) was >0.3 are asterisked, underlined when >0.4, and in bold when >0.5. Those for which η was <0.2 are enclosed in brackets.

Rank	Finger temperature	BVP amp	BVP amp	CEC
1	**Participant**	**Participant**	**Prior Rx (any)**	30
2	**Prior Rx (any)**	**Prior Rx (any)**	**Core temp**	20
3	**R tolerance**	Prior acupuncture	**Favourite**	16
4	L tolerance	Prior TENS	L OR R twitch	16
5	R threshold	Room temperature	Room temp	13
6	L threshold	L tolerance*	Gender	11
7	Room temperature*	Age	Handedness	11
8	L OR R twitch*	R tolerance	Stim Hz	8
9	L twitch*	Prior EA	L/R threshold	8

10	R twitch*	Slot order	L/R tolerance	4
11	Handedness*	L threshold		
12	Prior TENS*	Core temperature	other	
13	Favourite	Handedness	factors	
14	*[Age]*	Favourite	not	
15	*[Stimulation Hz]*	[L twitch]	analysed	
16	*[Slot order]*	[Gender]		
17	*[L/R tolerance (mean)]*	[L twitch]		
18	*[Core temperature]*	[Stimulation Hz]		
19	*[Gender]*	[R threshold]		
20	[Prior acupuncture]	[R twitch]		
21	[L/R threshold (mean)]	[L/R tolerance (mean)]		
22	[Prior EA]	[L/R threshold (mean)]		
Average	η = 0.309	η = 0.262		
Median	η = 0.311	η = 0.223		

Eta (η) values for the items in italics were ranked in a slightly different order for median rather than mean finger temperature, but remained within the 14 to 19 range of ranks. Given the few left-handers in the sample to date, the rankings for the Handedness factor for both finger temperature and BVP amplitude should not be taken too seriously. In addition to those factors listed for BVP amplitude, CEC was also high for Slot order.

CEC, calculated for the various BVP amplitude measures, was highest for the DIRECT measures, then the BASE, and least for the PRE-POST measures. As elsewhere in this presentation, the DIRECT measures appear to be the most useful

Eta (η) was also calculated for the effect of short slots on the data, and for the whole group this was <0.2 for all variants of finger temperature and BVP amplitude except for two of the BASE BVP amplitude measures. In contrast, it was >0.4 for 4 of the 9 PRV measures explored.

Overall, η was higher for finger temperature than for BVP amplitude.

Table 18 shows η for PTT (median, processed), computed for some major factors. Ranking was different if BASE or PRE-POST were used rather than PTT_DIRECT.

Table 18. *Eta* (η) for PTT (median, processed), computed for the main factors considered.

Rank	PTT
1	**Participant**
2	Time (slot order)
3	Prior Rx (any)
4	Stimulation Hz
5	[Favourite]
6	[Gender]
7	[Age]
Average	η = 0.245
Median	η = 0.216

Eta (η) for PTT_BASE and for PTT_PRE-POST was higher (median values 0.267 and 0.350, respectively).

Graphical illustrations and Tables for some of these factors follow. In each Figure, on the left are results for all slots, whether of the correct duration or short, and on the right the results when only slots of the correct duration were considered. Each graph shows mean and median group values of percentages normalised to the lowest mean or median for the three slots considered – stimulation, 5-minute post-stimulation, and the subsequent (10-minute post-stimulation) slot.

Prior treatment

Participants were asked whether they had prior experience of acupuncture, EA or TENS. Possible questionnaire responses were 'Yes' or 'No'. Responses to questions about the three modalities were all scored together.

A. Finger temperature (median, processed)

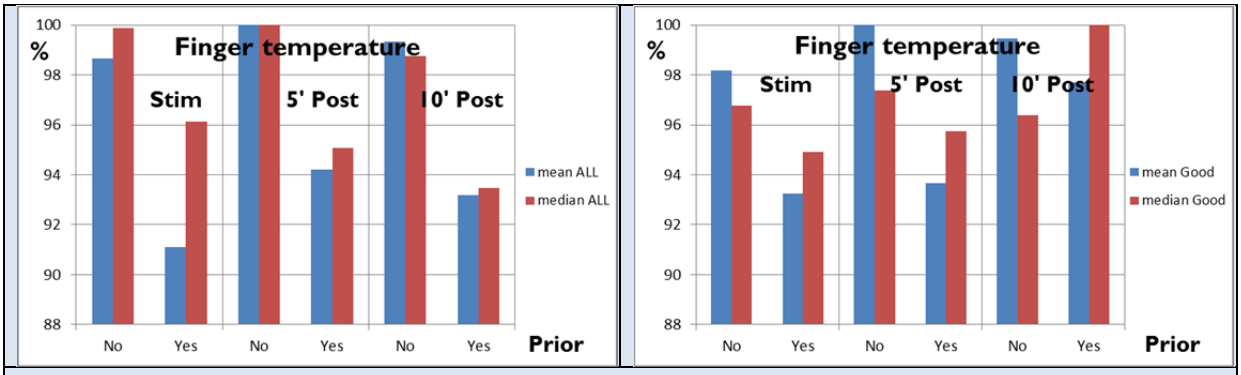


Fig 20. Finger temperature (median, processed), split according to Prior experience of related treatment. Group means and medians are shown in both parts, normalised to maximal values. **A.** For ALL slots. **B.** For Good slots only.

Note that within each of the three slots, the pattern is similar, with both means and median finger temperature greater in those who had *not* had prior experience of a related treatment. The only instance where this was not the case was for the median temperature in the 10-min Post-stimulation slots.

B. BVP amplitude (median, processed)

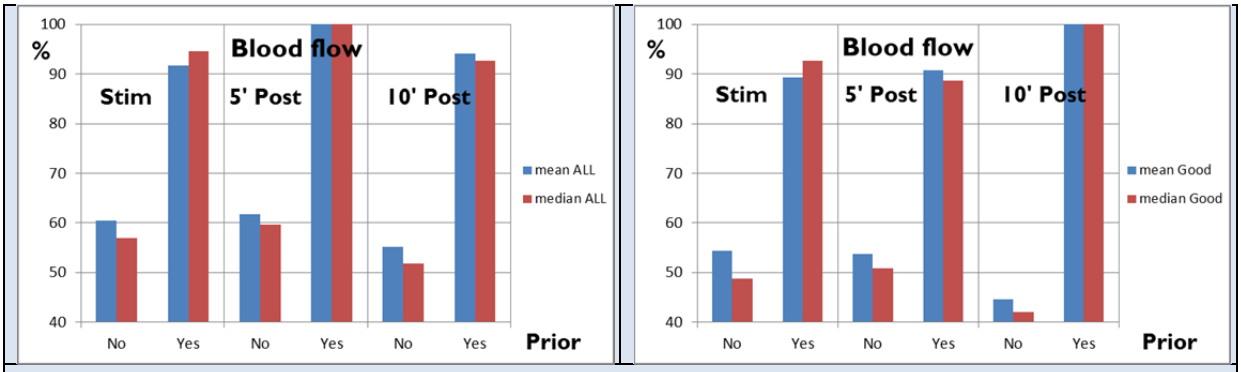


Fig 21. BVP amplitude (median, processed), split according to Prior experience of related treatment. Group means and medians are shown in both parts, normalised to maximal values. **A.** For ALL slots. **B.** For Good slots only.

Patterns of difference are similar for blood flow, regardless of slot duration and whether the mean or median was used. Thus this result appears more robust than the corresponding result for finger temperature (note too that the percentage differences shown here are much greater than those in **Fig 20**).

C. PTT (median, processed)

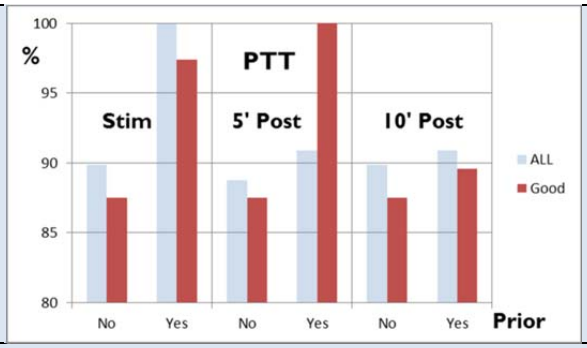


Fig 22. PTT (median, processed), split according to Prior experience of related treatment. Group medians are shown, normalised to maximal values.

In all three slots, PTT was greater for those who had prior experience of related treatment.

D. Significance of differences

The Mann-Whitney test was used (2-tailed, asymptotic significance), as the groups being compared were independent and not of the same size. Results are shown in **Table 19**.

Table 19. Significance of differences between factor levels for Prior experience of related treatment.

Dataset used	Median processed data	Stim	5-min Post-stim	10-min Post-stim
All slots (short & long)	Finger temp	n.s.	n.s.	n.s.
	BVP amplitude	p=0.001	p=0.001	p=0.001
Correct duration slots	Finger temp	n.s.	n.s.	n.s.
	BVP amplitude	p=0.001	p=0.003	p<0.001

During stimulation, the difference between factor levels was significant for PTT (p=0.04).

Favourite stimulation

Possible responses were 'Favourite'or 'Least favourite'. If both responses were given, the third stimulation was graded as 'medium'. If no 'Favourite' or 'Least favourite' was indicated, the response was scored as 0. This was also the case for the first four participants, who were not asked this question (suggested in conversation with one of the earlier participants).

A. Finger temperature (median, processed)

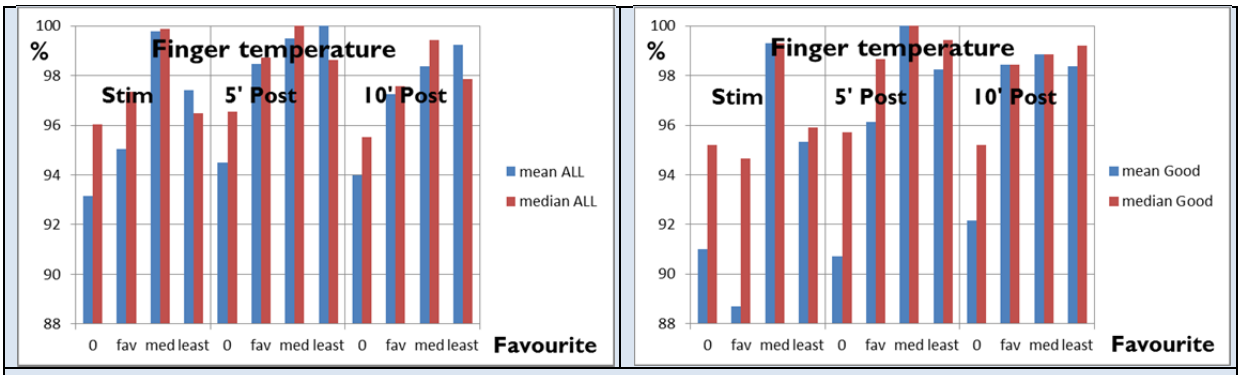


Fig 23. Finger temperature (median, processed), split according to Favourite stimulation. Group means and medians are shown in both parts, normalised to maximal values.
A. For ALL slots. **B.** For Good slots only.

Curiously, in most slots (except for ALL Stim mean, ALL 5-min Post-stimulation median and good 10-min Post-stimulation mean), finger temperature is lowest in response to the favourite stimulation, and higher for the least or medium favourite.

B. BVP amplitude (median, processed)

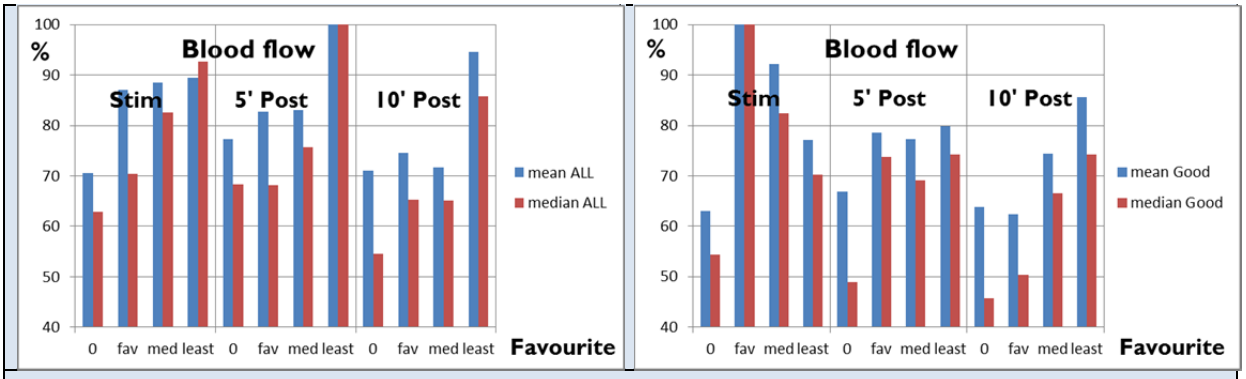


Fig 24. BVP amplitude (median, processed), split according to Favourite stimulation. Group means and medians are shown in both parts, normalised to maximal values. **A.** For ALL slots. **B.** For Good slots only.

In the Stimulation and 5-min Post-stimulation slots, BVP amplitude for ALL slots is lower in response to the favourite stimulation, higher for the others; in all three slots it is highest for the *least* favourite stimulation. However, if only the slots of correct duration are considered, the pattern is very different, with highest blood flow in response to the favourite stimulation in the Stimulation slot, but in response to the least favourite stimulation in the 10-min Post-stimulation slot.

As before, note that the percentage differences here are much greater than those for finger temperature.

C. PTT (median, processed)

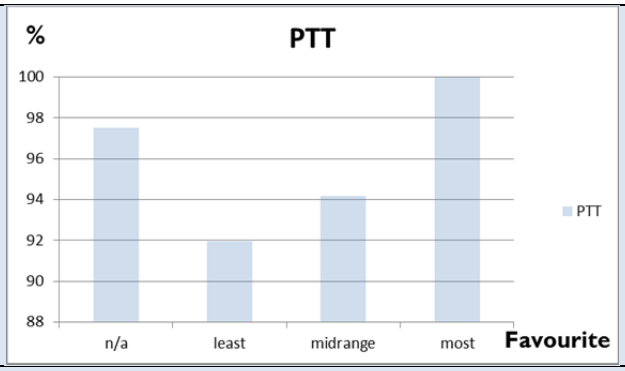


Fig 25. PTT (median, processed) during Stimulation, split according to Favourite stimulation. Group medians are shown, normalised to maximal values.

During stimulation, PTT was greatest for the favourite stimulation, and least for the least favourite.

D. Significance of differences

The Kruskal-Wallis H test was used (2-tailed, asymptotic significance), as the groups being compared were independent and not of the same size. Results are shown in **Table 20**.

Table 20. Significance of differences between factor levels for Prior experience of related treatment.

Dataset used	Median processed data	Stim	5-min Post-stim	10-min Post-stim
All slots (short & long)	Finger temp	n.s.	n.s.	n.s.
	BVP amplitude	n.s.	n.s.	n.s.
Correct duration slots	Finger temp	n.s.	n.s.	p=0.040
	BVP amplitude	n.s.	n.s.	n.s.

The Mann-Whitney test gave no results as significant when only 'favourite' and 'least favourite' were compared.

<u>The presence of finger or other hand muscle twitch</u>

Participants were asked if they noticed the presence of a twitch in either the left (L) or right (R) hands, or in both. If no twitch was experienced, the response was scored as 0. Not all participants were natural English speakers, and some may have understood 'twitch' to mean 'sensation'. It was not always clear if this was the case.

A. Finger temperature (median, processed)

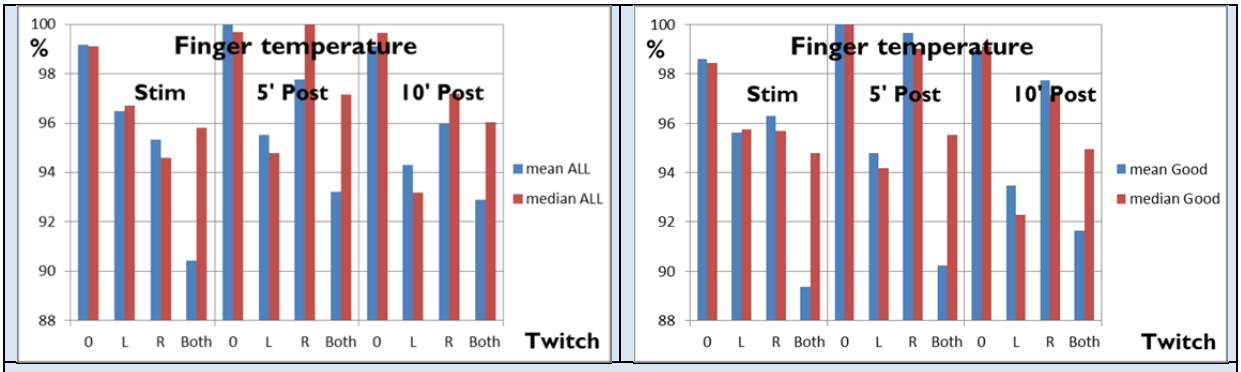


Fig 26. Finger temperature (median, processed), split according to presence of Twitch. Group means and medians are shown in both parts, normalised to maximal values. **A.** For ALL slots. **B.** For Good slots only.

In the Stimulation and 10-min Post-stimulation slots, the clearest pattern is that both mean and median temperature were highest without any noticeable twitch. In the 5-min Post-stimulation slot, lowest temperatures occur when twitch was experienced in the left hand, followed by bilateral twitch. Without twitch, or with a right-sided twitch, finger temperature was higher.

B. BVP amplitude (median, processed)

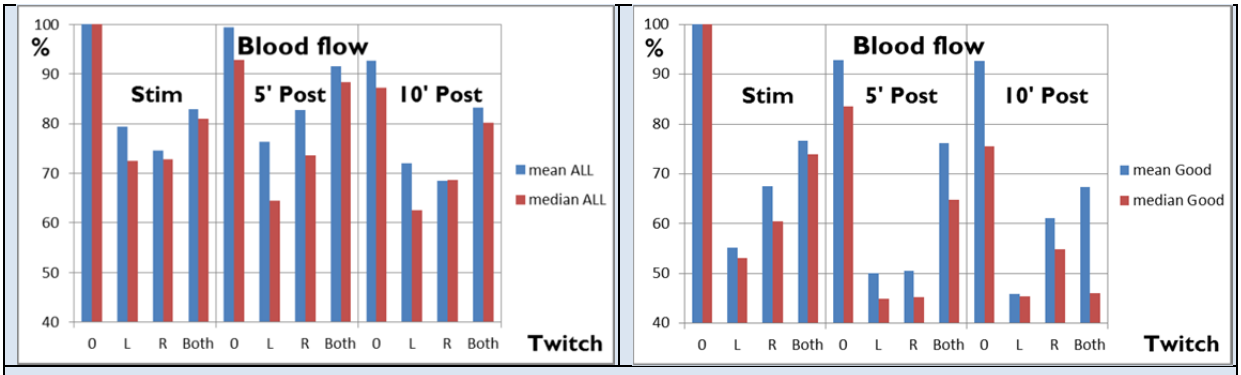


Fig 27. BVP amplitude (median, processed), split according to presence of Twitch. Group means and medians are shown in both parts, normalised to maximal values. **A.** For ALL slots. **B.** For Good slots only.

Blood flow was consistently greatest *without* muscle twitch, followed by bilateral twitch except for median blood flow in the 10-min Post-stimulation slot.

Again, the percentages here are much greater than those for finger temperature.

C. PTT (median, processed)

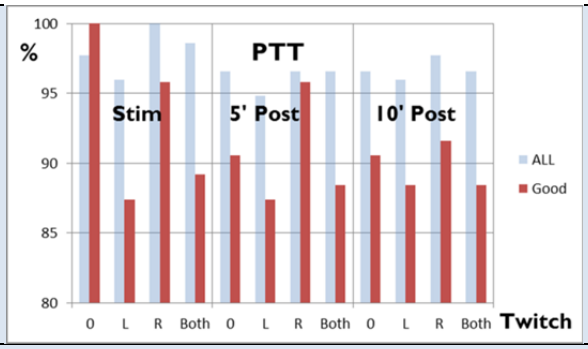


Fig 28. PTT (median, processed) during Stimulation, split according to presence of Twitch. Group medians are shown, normalised to maximal values.

Muscle twitch in the left hand only (L) results in lower PTT than no twitch. R twitch appears to enhance PTT more than bilateral twitch.

D. Significance of differences

The Kruskal-Wallis H test was used (2-tailed, asymptotic significance) because although the groups being compared were partially overlapping, they were not of the same size. Results are shown in **Table 21A.**

Table 21A. Significance of differences between all factor levels for experience of Twitch.				
Dataset used	**Median processed data**	**Stim**	**5-min Post-stim**	**10-min Post-stim**
All slots (short & long)	Finger temp	n.s.	n.s.	n.s.
	BVP amplitude	n.s.	n.s.	n.s.
Correct duration slots	Finger temp	n.s.	p=0.014	n.s.
	BVP amplitude	n.s.	n.s.	n.s.

The Mann-Whitney test was used for comparisons of two factor levels (**Table 21B**).

Table 21B. Significance of differences between two factor levels for experience of Twitch.				
Dataset used	**Median processed data**	**Stim**	**5-min Post-stim**	**10-min Post-stim**
All slots (short & long)	Finger temp	0.035[a]	0.025[a]	0.025[a]
	BVP amplitude	n.s.	n.s.	n.s.
Correct duration slots	Finger temp	0.032[a]	0.002[a]	p=0.021[a]
	BVP amplitude	n.s.	n.s.	p=0.036[a]; p=0.024[b]

a. Left or right vs no twitch; b. Left vs no twitch

For PTT during stimulation, there was a significant difference (p=0.041) between Left or right and no twitch for the raw, but not processed, data.

<u>Stimulation threshold ampltitude</u>

Participants were asked to state when they first felt the TEA stimulation, which was turned up first on the left hand, and then on the right. The 'dial unit' readings on the Equinox unit were recorded. In each session, they were classed as 'high', 'medium' and 'low' for the three stimulation slots (although for some participants threshold might be at only one or two settings, not three).

A. Finger temperature (median, processed)

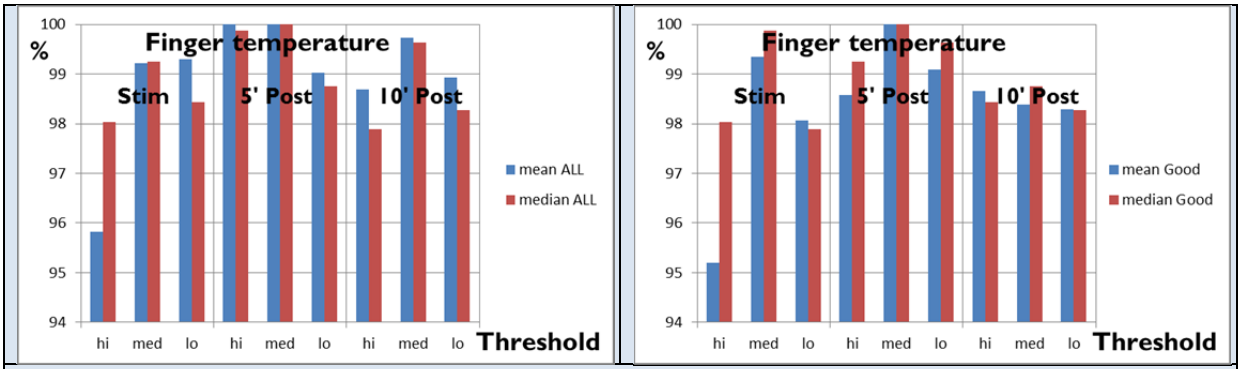


Fig 29. Finger temperature (median, processed), split according to stimulation Threshold amplitude. Group means and medians are shown in both parts, normalised to maximal values. **A.** For ALL slots. **B.** For Good slots only.

During stimulation, finger temperature was lowest when threshold was high. Post-stimulation, it was highest for a 'medium' threshold. Otherwise there was little consistency of response.

B. BVP amplitude (median, processed)

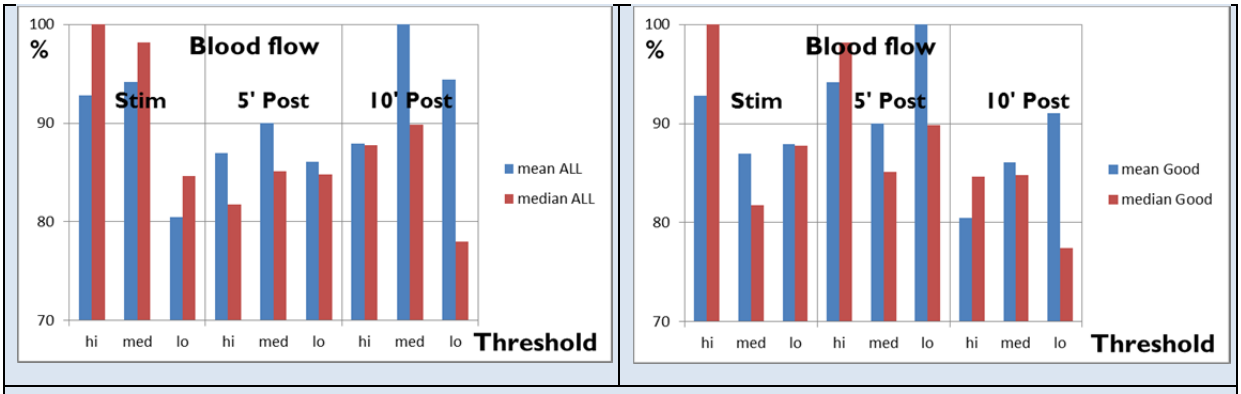


Fig 30. BVP amplitude (median, processed), split according to stimulation Threshold amplitude. Group means and medians are shown in both parts, normalised to maximal values.
A. For ALL slots. **B.** For Good slots only.

Median blood flow was highest for a high threshold in the Stimulation slot, but this was not the case for mean blood flow (for all slots, both short and of the correct duration). There was no clear pattern for blood flow in the 5-min Post-stimulation slot, but in the 10-min Post-stimulation slot mean blood flow was lowest for a high stimulation threshold, median blood flow lowest for a low stimulation threshold.

Again, the percentages here are somewhat greater than those for finger temperature.

C. PTT (median, processed)

Data was not examined for this factor.

D. Significance of differences

The Kruskal-Wallis H test was used (2-tailed, asymptotic significance) because the groups being compared were only partially overlapping and not of the same size. Results are shown in **Table 22**.

Table 22. Significance of differences between all factor levels for stimulation Threshold amplitude.

Dataset used	Median processed data	Stim	5-min Post-stim	10-min Post-stim
All slots (short & long)	Finger temp	n.s.	n.s.	n.s.
	BVP amplitude	n.s.	n.s.	n.s.
Correct duration slots	Finger temp	n.s.	n.s.	n.s.
	BVP amplitude	n.s.	n.s.	n.s.

The Mann-Whitney test gave no results as significant.

Stimulation tolerance amplitude

Once the threshold amplitude had been attained and recorded, stimulation level was increased until it was felt as 'strong' but still comfortable, and participants did not want it any higher. If it was then very quickly felt by the participant not to be strong enough, it was increased again. Otherwise, it was maintained at the same 'dial unit' setting throughout the five minutes of stimulation.

A. Finger temperature (median, processed)

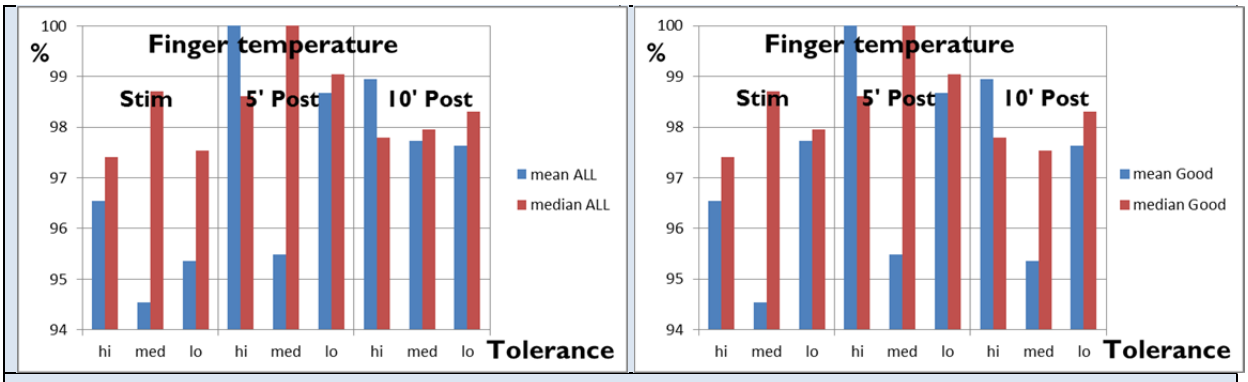


Fig 31. Finger temperature (median, processed), split according to stimulation Tolerance amplitude. Group means and medians are shown in both parts, normalised to maximal values. **A.** For ALL slots. **B.** For Good slots only.

During and after stimulation, median finger temperature was *least* for a high stimulation tolerance (apart from in the 10-min Post-stimulation slot).Following stimulation, mean finger temperature was *greatest* for a high stimulation tolerance. Otherwise there was little consistency among the results.

B. BVP amplitude (median, processed)

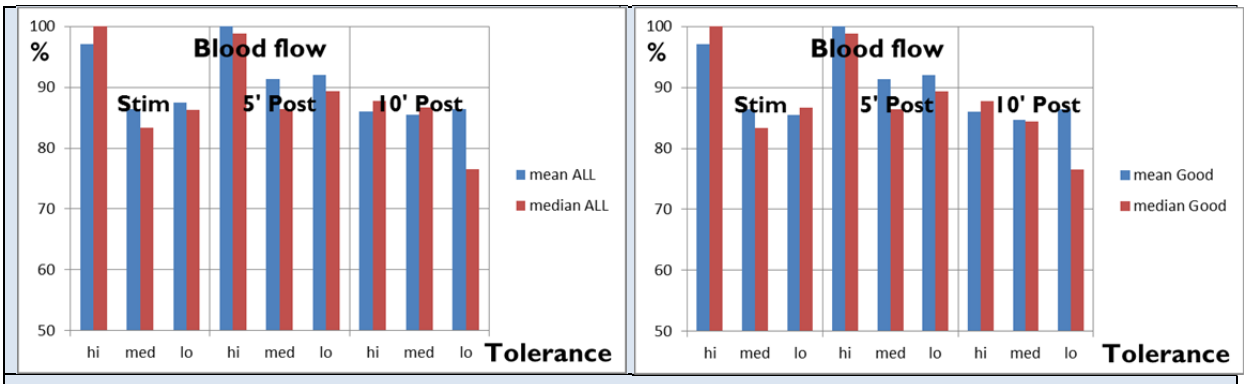


Fig 32. BVP amplitude (median, processed), split according to stimulation Tolerance amplitude. Group means and medians are shown in both parts, normalised to maximal values. **A.** For ALL slots. **B.** For Good slots only.

These results are larger and much more consistent than those for temperature. Median blood flow is consistently greatest for a high tolerance level, and the same is true for mean blood flow, except in the 10-min Post-stimulation slot.

Again, the percentages here are somewhat greater than those for finger temperature.

C. PTT (median, processed)

Data was not examined for this factor.

D. Significance of differences

The Kruskal-Wallis H test was used (2-tailed, asymptotic significance) because the groups being compared were only partially overlapping and not of the same size. Results are shown in **Table 23**.

Table 23. Significance of differences between all factor levels for stimulation Threshold amplitude.

Dataset used	Median processed data	Stim	5-min Post-stim	10-min Post-stim
All slots (short & long)	Finger temp	n.s.	n.s.	n.s.
	BVP amplitude	n.s.	n.s.	n.s.
Correct duration slots	Finger temp	n.s.	n.s.	n.s.
	BVP amplitude	n.s.	n.s.	n.s.

The Mann-Whitney test gave no results as significant.

Time (slot order)

Values of the three outcome measures were compared for the three stimulation slots, using only the medians, not the means. For skin temperature, values were lowest in slot 2 and highest in slot 5. For BVP amplitude, if only the slots of the correct duration were considered, values were highest in slot 2, lowest in slot 8, but for ALL the slots, BVP amplitude was highest in slot 8. For PTT, the pattern was similar to that for skin temperature.

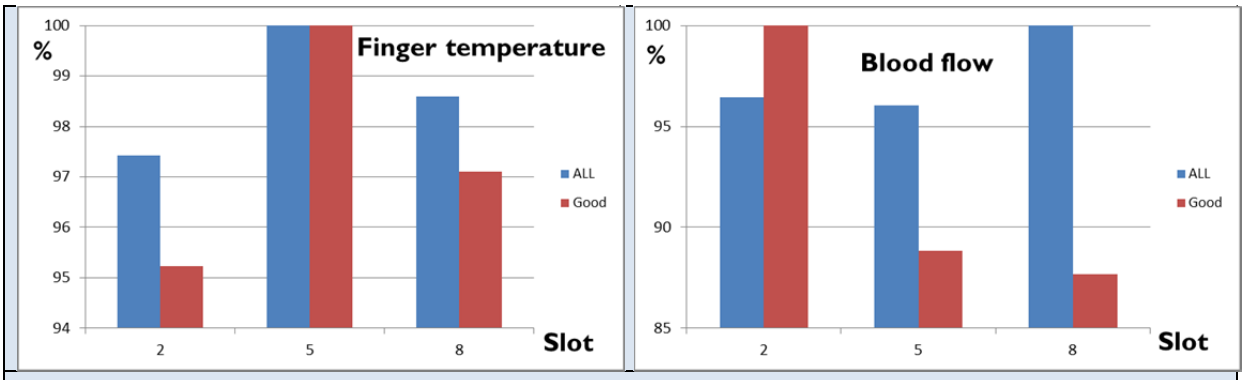


Fig 33. Outcome measures (median, processed) during stimulation, split according to order of stimulation and normalised to maximal values. **A.** Skin temperature. **B.** BVP amplitude.

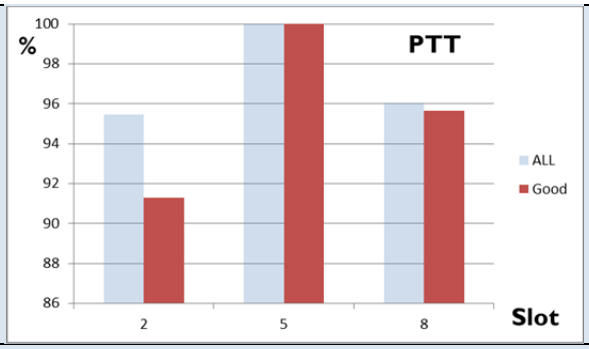


Fig 34. PTT (median, processed) during stimulation, split according to order of stimulation.

Significance of differences

Wilcoxon signed ranks test for paired samples was used (2-tailed, asymptotic significance), for ALL data during stimulation only. Results are shown in **Table 24**.

Table 24. Significance of differences between factor levels for Time (Slot order).

Dataset used	Median processed data	Slot 2 vs Slot 5	Slot 2 vs Slot 8	Slot 5 vs Slot 8
All slots (short & long)	Finger temp	p<0.001	p=0.011	p<0.001
	BVP amplitude	p<0.001	p<0.001	p<0.001
	PTT	p=0.009	n.s.	p=0.025

The effects of time (slot order) are clearly important (see too the sections on 'Changes over time', above). Those of stimulation frequency may be less so.

The effects of stimulation frequency on temperature, blood flow and PTT

In the above list, *eta* (η) is small for stimulation frequency (0.174 for mean and 0.176 for median finger temperature, ALL slots; 0.288 for median, Good slots), indicating that the factor of stimulation frequency is likely to have only a minor effect on this dependent variable.

A. Finger temperature (median, processed)

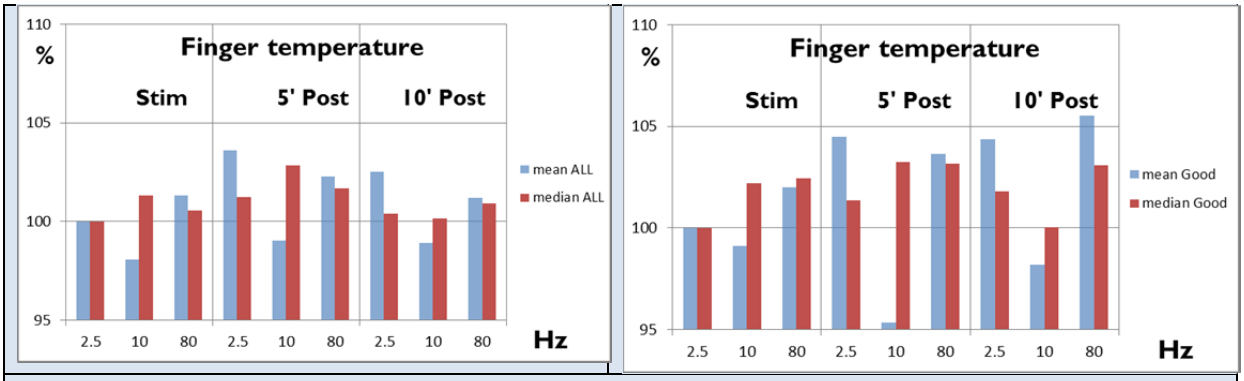


Fig 35. Finger temperature (median, processed), split according to stimulation Frequency. Group means and medians are shown in both parts, normalised to values for 2.5 Hz. **A.** For ALL slots. **B.** For Good slots only.

During Stimulation, highest mean finger temperature occurred at 80 Hz, but in the 5-min Post-stimulation slot in response to 2.5 Hz. Median finger temperature responded least during and immediately after stimulation at 2.5 Hz, but in the 10-min Post-stimulation slot at 10 Hz. Median finger temperature in the 5-min post-stimulation slot was greatest at 10 Hz, but least at 10 Hz in the 10-min Post-stimulation slot (and greatest at 80 Hz). It is difficult to find any consistency in these results.

Maximal and minimal finger temperature (median, processed) occurred at different frequencies for the 12 different participants for which data was recorded. Four showed maximal temperature both during and immediately after 2.5 Hz stimulation, two during and after 10 Hz, and three during and after 80 Hz stimulation. Two others showed maximal temperature during but not after 2.5 Hz stimulation, two others during but not after 10 Hz stimulation (and two after but not during 10 Hz stimulation), and one other after but not during 80 Hz stimulation.

B. BVP amplitude (median, processed)

As for finger temperature, *eta* (η) is small for stimulation frequency (0.120 for median BVP amplitude, ALL slots; 0.143 for Good slots), indicating that the factor of stimulation frequency is likely to have only a minor effect on blood flow as a dependent variable.

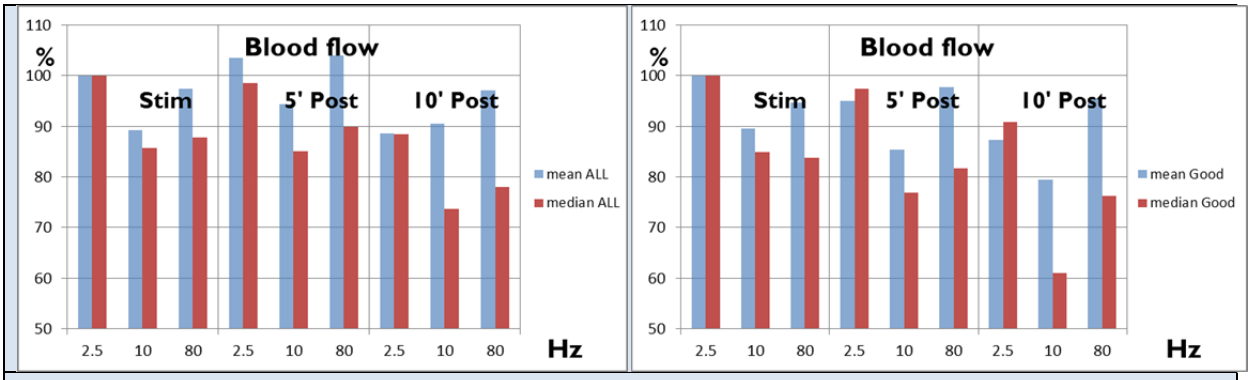


Fig 36. BVP amplitude (median, processed), split according to stimulation Frequency. Group means and medians are shown in both parts, normalised to values for 2.5 Hz. **A.** For ALL slots. **B.** For Good slots only.

Clearly, median blood flow increased most both during and following stimulation at 2.5 Hz, and following stimulation was lowest at 10 Hz. Mean blood flow during stimulation was also greatest at 2.5 Hz and least at 10 Hz, but following stimulation was greatest at 80 Hz, although again least at 10 Hz.

A first impression is also that blood flow *maxima* (as described above, pp 19-20) occurred more in response to 2.5 Hz stimulation (7 sessions) than 80 Hz (4 sessions) or 10 Hz (2 sessions). However, in 4 sessions it was difficult to determine which peak was maximal. Further cases would be required to confirm these impressions. In the meantime, individual blood flow responses to stimulation frequency in the present cohort were investigated, looking at maximal and minimal blood flow (median, processed) both during and following stimulation. The maxima and minima occurred at different frequencies for different participants, most clearly during (rather than after) stimulation. These results are shown in **Fig 37**.

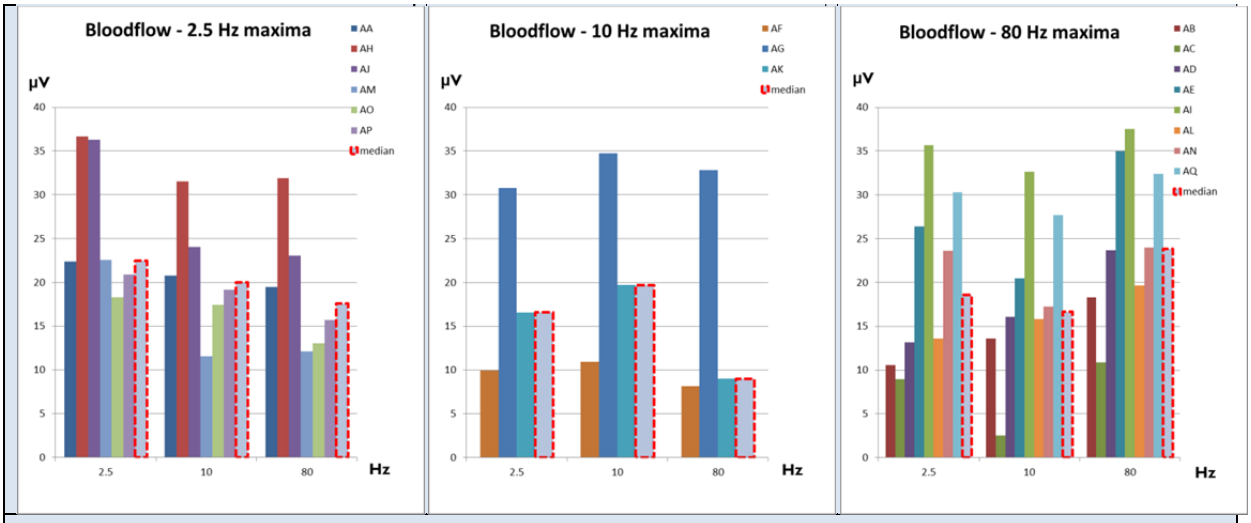


Fig 37. Maximal and minimal blood flow (median, processed) at different frequencies for different participants, shown during stimulation.

Whether the mean or median, raw or processed data was considered, six participants consistently showed BVP amplitude maxima at 2.5 Hz, six at 80 Hz, but only 3 at 10 Hz. Two other participants(Ab and AQ) were 'marginals', changing their allegiance depending on whether the data was raw or processed: Maximal flow for AB occurred in three blood flow measures at 80 Hz (in all but mean processed BVP amplitude), while for AQ flow was maximal for 2.5 Hz only in the raw data, being at 80 Hz in the processed data.

Of the six participants (50% of the sample) showing maximal temperature during 2.5 Hz stimulation, four also showed maximal BVP amplitude at 2.5 Hz. All three showing maximal temperature during 80 Hz stimulation also showed maximal BVP amplitude at that frequency, but only one of the four showing maximal temperature during 10 Hz also showed maximal BVP amplitude at that frequency.

C. PTT (median, processed)

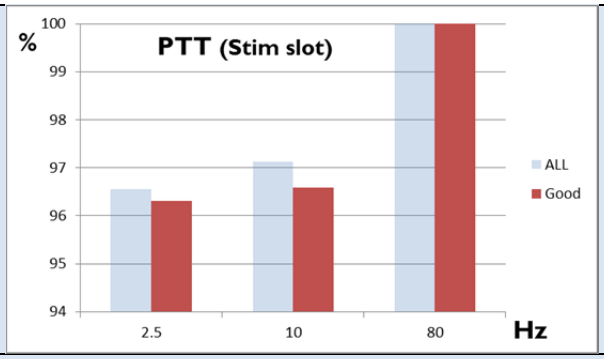


Fig 38. PTT (median, processed), split according to stimulation Frequency, for Good and ALL data, and normalised to maximum values.

80 Hz stimulation resulted in longer PTT than did the other two stimulation frequencies.

D. Significance of differences

The Friedman test (nonparametric equivalent of the 1-way analysis of variance, ANOVA, for multiple comparisons) comparing core (oral) temperature for all three stimulation frequencies showed significance (p=0.025). However, in both Wilcoxon and Binomial tests (see below for explanation), core temperature was significantly lower only at 80 Hz than at 2.5 Hz (p=0.017 for both tests).

Neither the Friedman nor Wilcoxon tests showed any significant differences in finger temperature (median, processed) for the stimulation frequencies.

When applied to the four DIRECT BVP amplitude measures, the Wilcoxon signed-ranks test did not demonstrate significance for any comparisons, and when the Binomial test was added to determine if the ratio of positive to negative differences in the measures for the two frequencies concerned was different to that expected by chance (0.5), results were significant (p=0.049) only for the mean DIRECT BVP amplitude (raw or processed). This suggests that apparent differences in blood flow at different stimulation frequencies may have more to do with large effects for particular individuals than being a genuine group effect. However, there were consistently more positive than negative differences for BVP amplitude at 2.5 Hz minus flow at 10 Hz for all four DIRECT BVP amplitude measures, suggesting as in **Fig 36** that indeed in this experimental situation **blood flow may be greater at 2.5 Hz than at 10 Hz**.

Of the other BVP amplitude variants, only mean_BASE showed significance (once more for **2.5 vs 10 Hz**) in the both the Wilcoxon (p=0.028) and Binomial (p=0.049) tests. Again, more differences were positive than negative (as they were for the other BASE measures of BVP amplitude, for which the Binomial test was not significant).

None of the BVP amplitude measures showed significant differences in the Friedman test for multiple comparisons of stimulation frequency.

Neither the Friedman nor Wilcoxon tests showed any significant differences in PTT (median, processed) for the stimulation frequencies.

Percentage differences for BVP amplitude (median, processed) when compared for the three stimulation frequencies were around 37% for the Good slots, reducing to around 19% for ALL data. For finger temperature, comparable differences were around 0.92% for ALL (2.22% Good), and for PTT only 0.60% for ALL (0.29% Good).

Other effects of stimulation frequency

In contrast to the somewhat equivocal results for BVP amplitude, a number of the *stimulation* amplitude parameters showed marked differences at different stimulation frequencies:

Left hand amplitude threshold (L.thresh)
Left hand amplitude tolerance (L.tol)
Right hand amplitude threshold (R.thresh)
Right hand amplitude tolerance (R.tol)
Mean of left and right hand amplitude thresholds (LR.thresh)
Mean of left and right hand amplitude tolerances (LR.tol)

In both Wilcoxon and Binomial tests, these all showed *significant* differences for all three pairs of stimulation frequencies, with amplitudes being higher at 10 Hz than at 2.5 or 80 Hz, and at 2.5 Hz than at 80 Hz: $Amp_{10Hz} > Amp_{2.5Hz} > Amp_{80Hz}$.

Thus stimulation amplitude is clearly an important confounder.

Stimulation frequency and subjective mood

Assessed using the NRS, there were similar changes in subjective mood at all three stimulation frequencies, as shown in the graph of mean values in **Fig 39** (a graph of *median* values shows little difference for any of the items except Relaxation, greatest at 10 Hz, and Good mood, least at 80 Hz).

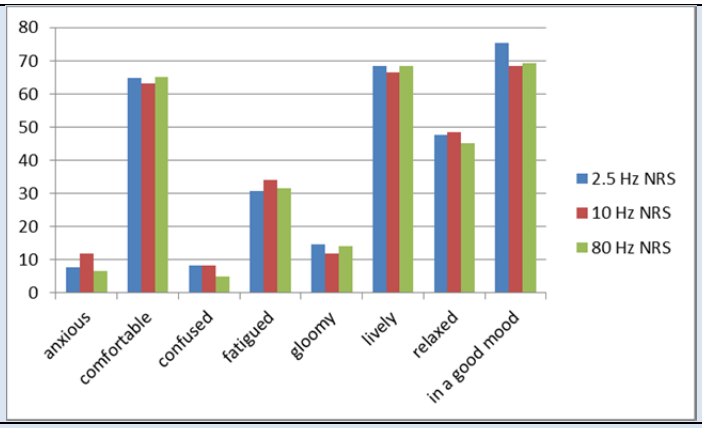


Fig 39. Mean values of the various NRS mood item scores following stimulation at the different frequencies.

Using two-tailed Wilcoxon and Marginal homogeneity tests, the only significant difference for any subjective mood item appeared to be for feeling 'in a good mood', with a better mood during or following stimulation at 2.5 Hz than at 80 Hz ($p=0.031$ and $p=0.037$, respectively).

At 2.5 and 80 Hz, the greatest differences in mood between participants were in fatigue, with those who showed greatest blood flow being more fatigued than the other participants; at 10 Hz, those who showed greatest blood flow were less relaxed than the other participants (compare **Fig 37** and also p 30 above).

Significant associations of subjective mood with temperature and blood flow at different frequencies

Significant associations (*rho*) are shown for the Stimulation slots, with those only for the 5-min Post-stimulation slots in square brackets. There did not appear to be any significant associations of subjective mood with blood flow [**Table 25**].

Table 25. Significant associations of subjective mood with temperature and blood flow at the different stimulation frequencies.

	Significant at 2.5 Hz	Significant at 10 Hz	Significant at 80 Hz
Core temperature	Relaxed 0.614		
Finger temperature	Anxious 0.644[ab] Comfortable -0.624 Confused 0.767[ab] Fatigued 0.790[ab] Gloomy 0.768[ab] Lively -0.668	[Anxious 0.593[c]] Fatigued 0.650[cb]	
BVP amplitude			

a. Smaller values in subsequent slot; b. Smaller values in next but one slot; c. Larger values in subsequent slot; d. Larger values in next but one slot.

It is curious that, as above, finger temperature is negatively correlated with positive items ('Comfortable' and 'Lively'), but positively with more negative moods.

Other interactions of interest

Eta (η) was calculated for various other interactions, suggesting for example that age, and to a lesser extent gender, affect stimulation amplitude thresholds and tolerances, and that Core temperature and Favourite stimulation may also vary with age.

A Chi^2 (χ^2) test showed no significant association between Favourite stimulation and frequency. In contrast, the Chi^2 (χ^2) test did show an association between prior treatment and Right stimulation threshold (p=0.039 for Prior TENS; p=0.018 for any Prior Rx), and between any prior Rx and the presence of twitch (p=0.010 for R twitch; p=0.048 for L OR R twitch).

Such interactions may need to be considered when building an explanatory statistical model.

Estimated sample sizes for further research

Sample size estimation on the basis of the results obtained so far were used to determine how many participants will be required to attain *statistically* significant results when comparing outcome measures for the different factor levels listed above. It should be noted that to estimate the sample size required for a *clinically* significant result would have necessitated using 'minimum important differences' from the literature, together with the standard deviations found in this study, not the actual group means *and* standard deviations we found [Neil Spencer. Personal communication. 25 Feb 2015].

Sample sizes (SS) were estimated in G*Power (version 3.1.9.2, 2014) [Faul et al. 1992-2014], using the parametric or nonparametric T-test family method appropriate for the data (difference between dependent or independent means, Wilcoxon signed-rank test, or Wilcoxon-Mann-Whitney test). Two-tailed tests were used, with power set at 0.80. Calculations were based on the data from ALL slots (of both short and correct duration) unless otherwise stated. Figures in parentheses show effect size as Cohen's d, dz (standardised d) or equivalent, depending on the test used. Sample sizes >100 are not realistically achievable in this series of studies.

1. Finger temperature (median, processed) – assessing the effects of stimulation frequency

For changes compared to baseline (BASE), lowest SS was found for **10 Hz** during stimulation ('stim'), following stimulation ('5-min post'), as well as in the following slot ('10-min post'). When DIRECT values were compared for the different frequencies, SS during stim slots was greatest for 10 Hz vs 80 Hz, least for 2.5 Hz vs 80 Hz during stimulation, but for 2.5 Hz vs 10 Hz following stimulation [**Table 26**].

Table 26. Sample size estimates for differentiating finger temperature according to stimulation frequency (ALL data).

F temp	base vs 10	base vs 2.5	base vs 80	10 vs 2.5	2.5 vs 80	80 vs 10
stim slot	94 (0.300)	>100	88 (0.309)	>100	>100	>100*
5-min post	41 (0.460)	44 (0.243)	61 (0.374)	**63** (0.359)	>100	>100
10-min post	>100	65 (0.363)	83 (0.320)	**46** (0.423)	>100	>100

* This reduced to **52** if only slots of the correct duration were used.

Change from baseline appears to be greater following than during stimulation, but it is not possible to say which stimulation frequency evokes most or least change. It may be more possible to find a difference in finger temperature for 10 vs 2.5 Hz than for the other stimulation frequency pairs, but this is more likely following than during stimulation.

2. Blood flow – assessing the effects of stimulation frequency

A. BVP amplitude (median, processed) [**Table 27**]

Table 27. Sample size estimates for differentiating blood flow (BVP amplitude) according to stimulation frequency (ALL data).

BVP amp	base vs 10	base vs 2.5	base vs 80	10 vs 2.5	2.5 vs 80	80 vs 10
stim slot	15 (0.808)	58 (0.384)	58 (0.384)	**36** (0.485)	>100	89 (0.301)
5-min post	23 (0.640)	>100	>100	>100	>100	**54** (0.392)
10-min post	14 (0.870)	23 (0.628)	39 (0.477)	>100	>100	83 (0.313)

Change from baseline appears to be more detectable when using 10 Hz than the other stimulation frequencies, and is more likely in the '10-min post' slots than during stimulation.

Note that here there is greater differentiation between 2.5 and 10 Hz effects *during* than *after* stimulation, but immediately after rather than during stimulation for 10 Hz vs 80 Hz.

B. MMDiff [**Table 28**]

Table 28A. Sample size estimates for differentiating blood flow (MMDiff) according to stimulation frequency (ALL data).

MMDiff	base vs 10	base vs 2.5	base vs 80	10 vs 2.5	2.5 vs 80	80 vs 10
stim slot	50 (0.407)	67 (0.350)	>100	>100	>100	70 (0.341)
5-min post	81 (0.316)	>100	>100	>100	>100	85 (0.308)
10-min post	61 (0.365)	29 (0.548)	>100	>100	>100	>100

Change from baseline appears to be more detectable when using 10 Hz than the other stimulation frequencies, and is least likely in the post-stimulation than the other slots.

If only comparisons between the correct duration slots are considered (and excluding comparisons listwise, i.e. when there are missing values for any variable in the comparison), SS may be more manageable [**Table 28B**].

Table 28B. Sample size estimates for differentiating blood flow (MMDiff) according to stimulation frequency (Good data).

MMDiff	base vs 10	base vs 2.5	base vs 80	10 vs 2.5	2.5 vs 80	80 vs 10
stim slot				80 (0.325)	>100	**19** (0.698)
5-min post				**19** (0.705)	>100	**19** (0.716)
10-min post						

3. BVP amplitude, skin temperature and PTT (median, processed) – assessing the effects of time (slot order).

The effects of time were assessed for the whole session (slot 10 vs slot 1), for differences between the stimulation slots (2, 5 and 8), and so forth [**Table 29**].

Table 29. Sample size estimates for differentiating blood flow (BVP amp) according to time (slot order) (Good data).

	Slots 1 – 10	Stim slots (2, 5, 8)	5-min Post 3 – 9	5-min Post 3 – 6	5-min Post 6 – 9	10-min Post 4 – 7	10-min Post 4 – 10	10-min Post 7 – 10
SS (*dz*)	17 (0.751)	>100	25 (0.599)	>100	49 (0.419)	>100	23 (0.630)	42 (0.459)

As is already clear, time has a major impact on blood flow.

For finger temperature, only slots 2 vs 5 and 2 vs 8 showed achievable sample sizes (90 and 92, with effect sizes 0.306 and 0.303, respectively_.

For PTT, sample sizes were as in **Table 30.**

Table 30. Sample size estimates for differentiating PTT according to time (slot order) (Good data).

	Slots 1 – 10	Stim slots (2, 5, 8)	5-min Post 3 – 9	5-min Post 3 – 6	5-min Post 6 – 9	10-min Post 4 – 7	10-min Post 4 – 10	10-min Post 7 – 10
SS (dz)	>100	2-5 & 5-8 95 (0.297) 2-8 42 (0.458)	85 (0.316)	>100	>100	>100	>100	>100

4. BVP amplitude (median, processed) – assessing the effects of prior experience, stimulation preference and presence of twitch.

A summary of sample size estimates for these three factors is shown in **Table 31**.

Table 31. Sample size estimates for differentiating blood flow (BVP amp) according to Prior experience, Favourite stimulation and experience of Twitch (Good data).

	Prior (all)[a]	Favourite[b]	Twitch[c]
stim slot	32	>100	>100
5-min post	24	98	>100
10-min post	28	78	>100

a. Yes or No; b. Favourite vs least favourite; c. All comparisons

Thus the 'Slot order' factor demonstrates an effect on a par with that for 'Prior experience', for which prior MA had the strongest effect, and prior EA the weakest (although the numbers of those with prior experience of these modalities were small, so estimates may not be particularly accurate). 'Stimulation frequency' showed a smaller effect still, although sample sizes are still achievable. Preference had little effect. To achieve significance, the presence of a Twitch would require a larger group than we can recruit.

Discussion and conclusions

Collecting and analysing physiological time series data can be fraught with pitfalls, as this small study demonstrates. However, provided you keep your head, it is usually possible to salvage something useful even in the most seemingly inimical circumstances.

Here we learned that DIRECT measures of finger temperature and blood flow are more likely to be useful than BASE or PRE-POST, that missing (temperature) data and data corruption could be allowed for in analysis (even without embarking on more advanced statistical procedures), and that careful application of nonparametric methods to physiological time series data is probably going to provide more realistic estimates of outcome than rigidly attempting to use T-tests and Pearson's R.

In the previous poster in this series [Steffert & Mayor 2014], we used nSD (normalised SD, or the coefficient of variance), effect size (Cohen's *d*) and correlation ratio *eta* (η) to gauge the relative importance of various experimental design factors. Here we used η, effect size and sample size (SS) estimation to similar ends. SS estimation was required in order to assess how many more participants will have to be enrolled in our study to make it likely that we attain statistically significant results.

PTT has been investigated in hundreds of studies since 1980, but to our knowledge not previously in acupuncture-related research. It may be a potentially interesting outcome measure for use in this field because it reflects alterations in blood pressure, cardiac output and arterial stiffness [Kounalakis & Geladas 2009], with shortened PTT indicating sympathetically mediated increases in myocardial performance, vasoconstriction, or a combination of the two. Thus it has been used to evaluate the effects of short-term stress on cardiovascular function [Pan & Li 2007], as well as the longer-term changes associated with ageing, arterial stiffness (as in arteriosclerosis) and diabetes [Naschitz et al. 2004]. In future acupuncture-related research on HRV/PRV, PTT variability (PTTV) could also be investigated [Ma & Zhang 2006].

Specific findings

- A close correlation between two different methods of assessing blood flow (BVP amplitude and MMDiff), but no consistent correlation between either of these and PTT.
- Peaks in blood flow tended to occur in the five minutes after TEA. Peaks in finger temperature and blood flow often occurred together.
- Overall, as might be expected, there is a significant positive association between blood flow and temperature. However, although this is the case for most participants, for at least two the association was negative (although not significant). A negative correlation (significant, but not large) was found between finger temperature and PTT.
- A strong effect of time was observed, with both finger temperature and blood flow decreasing for most participants over the course of their session – presumably due to prolonged inactivity. Core (oral) temperature also showed a tendency to decrease with time (or remained steady), although room temperature remained steady or *in*creased with time. In contrast, despite considerable fluctuations over time, there was negligible difference in the median group PTT at the end of the session from that at baseline.
- Over time, there were also significant changes in subjective mood: significantly reduced POMS-SF tension/anxiety, vigour/activity and confusion/bewilderment, and significantly reduced NRS anxiety but also relaxation, and a significant increase in feeling comfortable.

The strong effect of time on temperature, blood flow and mood indicates the need for a control group to be incorporated in the next stage of this study, with the same or other participants exposed to either no or minimal intervention.

- Significant correlations between blood flow and time or frequency domain PRV measures suggest that PRV increases as blood flow decreases, and that that may be the case for nonlinear measures such as SampEn and D_2 as well.

A tentative interpretation of these somewhat confusing results would be that the *reduction* in skin blood flow and skin temperature that occurs during most sessions may be associated with increased parasympathetic tone and/or reduced sympathetic modulation, so a lower level of anxiety.

However, this contradicts the usual understanding that cutaneous vasodilatation is to some extent parasympathetically mediated (by cholinergic [Kálmán et al. 2002] or nitrergic [Toda & Okamura 2014] mechanisms), with cutaneous vasoconstriction being sympathetically mediated (by noradrenalin or endothelin [Burnstock & Ralevic 1994]), particularly in the extremities, where the fingertip vascular beds, for instance, are rich in sympathetic innervation [Krasnikov et al. 2014]. Thus, in the 'fight or flight' reflex, blood flow is preferentially diverted away from the skin to the lungs and skeletal muscle [Wikipedia 2015]. However, this may not be a simple either/or matter: at rest, as in this study, sympathetic pathways to the skin may also be *tonically* active [Gibbins 2013], and in this context it is possible that tonic sympathetic activation may have overridden the cutaneous phasic (parasympathetic) relaxation effect.

Further literature research needs to be carried out on the autonomic regulation of finger temperature, blood flow and PTT, and their interactions, in order to disentangle the sometimes apparently contradictory effects on them of stimulation (or even of simply sitting quietly for an extended period).

- Of the various possible factors explored, Prior experience (of acupuncture, EA or TEA) had greatest effect on BVP amplitude, and the presence of muscle twitch in the hand on finger temperature. Stimulation threshold and tolerance amplitudes did not have a significant effect on either.

Interestingly, prior experience of acupuncture was found in one study to enhance the increase in chorioretinal blood flow induced by manual acupuncture at LI4 (*hegu*) [Naruse et al. 2000]. Phasic muscle contractions have been considered an important factor in some studies of TENS and blood flow [Johnson 2014], and unpublished experimental research and systematic review support this conclusion [Tim Watson, personal communication, 23 Feb 2015.

- The factor of particular interest, stimulation frequency, had small and inconsistent effects on finger temperature. In contrast, 2.5 Hz appeared to result in greater blood flow (BVP amplitude) than 10 Hz or 80 Hz, although this was not significant. Six participants showed maximal blood flow at 2.5 Hz, and a further six at 80 Hz, but only three at 10 Hz. Individuality of responses requires further exploration.
- In contrast to stimulation frequency, there were significant differences between all three pairs of threshold and tolerance amplitudes (left, right, bilateral), with $Amp_{10Hz} > Amp_{2.5Hz} > Amp_{80Hz}$.
- Sample size (SS) estimates for significant finding suggest that it may be more possible to find a difference in finger temperature for 10 vs 2.5 Hz (SS 63, 46 following stimulation; effect size in the region of 0.36 to 0.42) than for the other stimulation frequency pairs, but that this is more likely following than during stimulation.
- For BVP amplitude, again a difference is most likely for 10 vs 2.5 Hz (SS 36 during stimulation; effect size 0.485). Taking only slots of the correct duration into account, for MMDiff the sample size required is smaller still (SS 19 post-stimulation; effect size 0.705).
- However, sample and effect sizes for the effect of time on BVP amplitude were lowest of all (e.g., for a difference between the first and last slot of sessions, SS 17; effect size 0.751).

Limitations and adverse effects

- Because of its internal design, an accurate estimate of the current output of the Equinox stimulator during use was not possible (although it had previously been calibrated using a standard resistive load).
- Combining the data capture methods described here with EEG recording may have skewed results. Two participants (both older women therapists rather than young students or academics) became quite uncomfortable during their sessions, with some subsequent short-lived emotional disturbance as well. They did not otherwise consider themselves as claustrophobic.

- The finger to which the BVP sensor was attached was not consistently the same for all participants. Although this should not have a major impact within-participant changes, the effects of using the sensor on different fingers should be investigated and taken into account in further investigations if it is not possible to use the same finger consistently for data collection.
- In at least one session, room temperature was very low initially, so that increasing temperature over time may have affected outcome.
- Whether cutaneous vasodilatation or vasoconstriction occurs may be a homoeostatic response dependent on baseline blood flow [Karita & Izumi 1995]. Further analysis is required to investigate this factor in the present data, which appear to indicate a counterintuitive decrease in skin blood flow with greater relaxation. The effect of participant respiration patterns may also be important, and our recordings will need to be examined.
- Some questions were not asked of the early participants, such as which was their favourite (or least favourite) stimulation, and whether they experienced muscle twitching during stimulation.
- Data analysis was complicated by the technical problems encountered.
- For example, because of a low sampling rate, PTT error level was high compared with actual data values.
- Furthermore, peak-to-peak PPT, like BVP amplitude, is quite dependent on finger temperature (local cold exposure) [Jaryal et al. 2009; Zhang & Zhang 2005]. Both these potential confounders of PTT should be taken into account in further investigations
- In this study, stimulation was only investigated for its effect on blood flow quite local to the stimulating electrodes. Further investigations should also examine the non-local effects of TEA, and also the effects of TEAS applied at points other than LI4 (*hegu*).
- The washout issue will have to be explored rigorously, possibly by incorporating a control group in the next stage of this study, with the same or other participants exposed to either no or minimal intervention.
- Calculations for *statistically* significant sample size estimation were carried out using actual group means and standard deviations. To estimate sample sizes required for *clinically* significant results will necessitate locating 'minimum important differences' (MIDs) from the literature ('To demonstrate that the treatment difference is statistically significantly larger than the MID, the lower bound of the confidence interval for the treatment difference should exceed the chosen MID' [Wyrwich 2011]).
- No statistical corrections were made for multiple comparisons (the 'fallacy of data dredging').
- More advanced statistical analysis (linear or multilevel modelling) will be required to examine the contribution of confounding factors such as stimulation amplitude.
- **Therefore all results can only be tentative, and will require confirmation using the larger sample planned, which will involve a more robust protocol and data collection**.

Erratum to poster presented at ARRC 2014

In this poster presentation, it was stated that 'The sum of η^2 for all factors considered was 0.678, suggesting that >2/3 of factors responsible for variance in outcomes were identified'. Given the methods used, this conclusion was unwarranted.

Attributions statement

DM: Study coordination and recruitment, NRS design, Kubios HRV, Excel and SPSS processing, G*Power sample size estimation, compilation of results and authoring this report
TS: Expert advice and training, Matlab processing, recording EEG and NeXus-10 data in the 17th session
RB: Recruitment; recording data in the first 16 sessions.

Acknowledgements

From the University of Hertfordshire, Professor Tim Watson (School of Health and Social Work) and Dr Na Helian (School of Computer Science) for supervisory advice; Neil Spencer (Director, Statistical Services and Consultancy Unit) and Alla Mashanova (Health and Human Sciences Research Institute) for their gentle statistical suggestions (which there was not always time to follow). Also: Jeroen Manten (Mind Media BV), for assistance with endless BioTrace+ and NeXus-10 queries; Zusana Kovacova (Herts EEG Biofeedback) for rescuing the study with a replacement temperature sensor; Aleksander Valjamae (Decision, Emotion and Perception lab, Department of Behavioural Sciences and Learning, Linkoping University, Sweden) for assistance with Matlab pulse transit time processing, and of course our spouses/partners, and the participants in our study, without whom none of this would have been possible.

Conflicts of Interests

It is always difficult to remain unbiased in an unblinded study. We hope that experimenter expectations (e.g. about the effects of stimulation frequency or amplitude and any resulting muscle twitch) have not inadvertently influenced outcomes.

References

Abu-Bader SH. 2010. Advanced and Multivariate Statistical Methods for Social Science Research. Lyceum Books, Chicago

Billauer E. 2012. peakdet: Peak detection using MATLAB. Available at: http://www.billauer.co.il/peakdet.html

Burnstock G, Ralevic V. 1994. New insights into the local regulation of blood flow by perivascular nerves and endothelium. British Journal of Plastic Surgery. Dec;47(8):527-543

Cohen S, Williamson G. 1988. Perceived stress in a probability sample of the United States. In: Spacapam S, Oskamp S. (eds.). The Social Psychology of Health: Claremont Symposium on applied psychology. Newbury Park, CA: Sage, pp. 31-67

Faul F, Buchner A, Erdfelder E, Lang AG. 1992-2014. G*Power Version 3.1.9.2. Available at: http://www.softpedia.com/get/Science-CAD/G-Power.shtml

Gibbins I. 2013. Functional organization of autonomic neural pathways. Organogenesis. Jul-Sep;9(3):169-175

Jaiswal M, Urbina EM, Wadwa RP, Talton JW, D'Agostino RB Jr, Hamman RF, Fingerlin TE, Daniels SR, Marcovina SM, Dolan LM, Dabelea D. 2013. Reduced heart rate variability is associated with increased arterial stiffness in youth with type 1 diabetes: the SEARCH CVD study. Diabetes Care. Aug;36(8):2351-2358

Jaryal AK, Selvaraj N, Santhosh J, Anand S, Deepak KK. 2009. Monitoring of cardiovascular reactivity to cold stress using digital volume pulse characteristics in health and diabetes. Journal of Clinical Monitoring and Computing. Apr;23(2):123-130

Johnson MI. 2014. Transcutaneous Electrical Nerve Stimulation (TENS). Research to support clinical practice. Oxford University Press, Oxford

Kálmán J, Szakács R, Török T, Rózsa Z, Barzó P, Rudas L, Papp JG, Janka Z. 2002. Decreased cutaneous vasodilatation to isometric handgrip exercise in Alzheimer's disease. International Journal of Geriatric Psychiatry. 2002;17(4):371-374

Karita K, Izumi H. 1995. Effect of baseline vascular tone on vasomotor responses in cat lip. Journal of Physiology. Feb 1;482 (Pt 3):679-685

Kistler A, Mariauzouls C, von Berlepsch K. 1998. Fingertip temperature as an indicator for sympathetic responses. International Journal of Psychophysiology. Jun;29(1):35-41.

Kounalakis SN, Geladas ND. 2009. The role of pulse transit time as an index of arterial stiffness during exercise. Cardiovascular Engineering. Sep;9(3):92-97

Krasnikov GV, Tiurina MĬ, Tankanag AV, Piskunova GM, Cheremis NK. 2014. [The significance of sympathovagal balance in the forming of respiration-dependent oscillations in cardiovascular system in human]. Fiziologiia Cheloveka. Jan-Feb;40(1):68-75

Ma HT, Zhang YT. 2006. Spectral analysis of pulse transit time variability and its coherence with other cardiovascular variabilities. Proceedings of the 28th IEEE EMBS Annual International Conference. New York, USA, Aug 30-Sept 3; 1:6442-6445

Malik M (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology) 1996.. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. European Heart Journal. 17(3): 354-381

Mayor DF. (Ed.). 2007. Electroacupuncture. A practical manual and resource. Churchill Livingstone, Edinburgh

Mayor D, Steffert T. 2013a. EEG and eyeblink response to different acupuncture modalities: preliminary results from four pilot studies. 15th ARRC Symposium, London South Bank University, 2 March 2013. Available at: http://www.qeeg.co.uk/electroacupuncture/eablink.htm

Mayor D, Steffert T. 2013b. Expectation and experience of the 'nonspecific' effects of acupuncture: Developing and piloting a set of questionnaires. BMAS/PMAS Meeting, 27-29 Sept 2013, Porto, Portugal. Available at: http://www.qeeg.co.uk/electroacupuncture/eaq1.htm

Naruse S, Mori K, Kurihara M, Nakajima N, Matsumoto Y, Kinoshita S, Ohyama Y, Nakamura T. 2000. [Chorioretinal blood flow changes following acupuncture between thumb and forefinger]. Nihon Ganka Gakkai Zasshi. Oct;104(10):717-723

Naschitz JE, Bezobchuk S, Mussafia-Priselac R, Sundick S, Dreyfuss D, Khorshidi I, Karidis A, Manor H, Nagar M, Peck ER, Peck S, Storch S, Rosner I, Gaitini L. 2004. Pulse transit time by R-wave-gated infrared photoplethysmography: review of the literature and personal experience. Journal of Clinical Monitoring and Computing. Dec;18(5-6):333-342

Pan RL, Li JK. 2007. A noninvasive parametric evaluation of stress effects on global cardiovascular function. Cardiovascular Engineering. Jun;7(2):74-80

Shacham S. 1983. A shortened version of the Profile of Mood States. Journal of Personality Assessment. 47(3):305-306

Steffert T, Mayor D. 2013. Does the cortical response to electroacupuncture depend on stimulation frequency? Results of a pilot EEG study first proposed at the AACP Conference in 2001. AACP Annual Conference, Wokefield Park, Reading, 18-19 May 2013. Available at: http://www.qeeg.co.uk/electroacupuncture/eaffr.htm

Steffert T, Mayor D. 2014. The fickleness of data: Estimating the effects of different aspects of acupuncture treatment on heart rate variability (HRV). Initial findings from three pilot studies. 16th Internationakl Acupuncture Research Symposium, King's College London, 29 March 2014. Available at: http://www.qeeg.co.uk/electroacupuncture/hrv1.htm

Terry PC, Lane AM, Lane HJ, Keohane L. 1999. Development and validation of a mood measure for adolescents. Journal of Sports Sciences. Nov;17(11):861-872

Terry PC, Lane AM, Fogarty GJ. 2003. Construct validity of the Profile of Mood States —Adolescents for use with adults. Psychology of Sport and Exercise. 4(2): 125–139

Toda N, Okamura T. 2014. Recent advances in research on nitrergic nerve-mediated vasodilatation. Pflugers Archiv: European Journal of Physiology. Oct 23. [Epub ahead of print]

Wikipedia. 2015. Autonomic nervous system. Available at: http://en.wikipedia.org/wiki/Autonomic_nervous_system

Wyrwich K. 2011. Statistical Considerations and PRO Interpretation Comparisons. ISPOR 16th Annual International Meeting. Baltimore, MD, 21-25 May. Available at: http://www.ispor.org/meetings/baltimore0511/presentations/WS17_Wyrwich.pdf

Zhang XY, Zhang YT. 2005. The effect of local cold exposure on pulse transit time. Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference. Shanghai, China, Sept 1-4; 4:3522-3525

APPENDICES

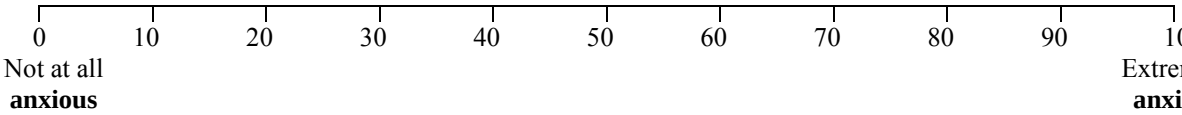
I. Exclusion criteria

People who have suffered past serious head injury, suffer from epilepsy or diabetes, currently have cancer or a respiratory condition that impairs nose breathing, or wear an implanted electronic device. Anyone who has a condition in which peripheral circulation is impaired (such as Raynaud's syndrome), or who currently has any shoulder, arm or hand injury. Those with severe physical or mental conditions or learning disabilities who would be unlikely to be able to complete their involvement in or otherwise comply with the requirements of the study. Those who are dependent on prescribed or other psychoactive substances, or who are very heavy users of caffeine, nicotine or alcohol, i.e. with a regular daily intake of more than 400 mg caffeine (equivalent to 4 cups of real filtered coffee, 10 cans of cola or 2 'energy shot' drinks), 20 cigarettes or equivalent, or two alcoholic drinks. Those who are currently undergoing other non-routine (i.e. not 'preventive') nonpharmacological or complementary medical treatments (depending on circumstances).

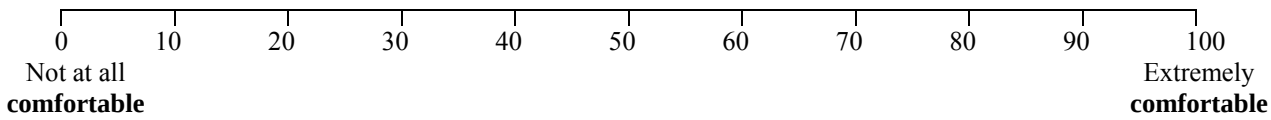
II. Composite NRS for current relaxation and mood

Place a cross on each line to represent **how you feel right now***.

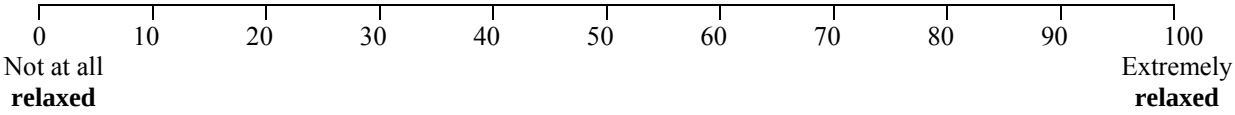
‘anxious’



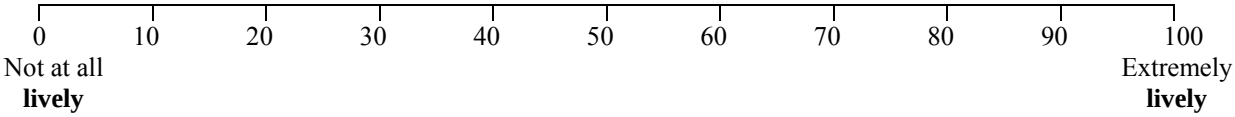
‘comfortable



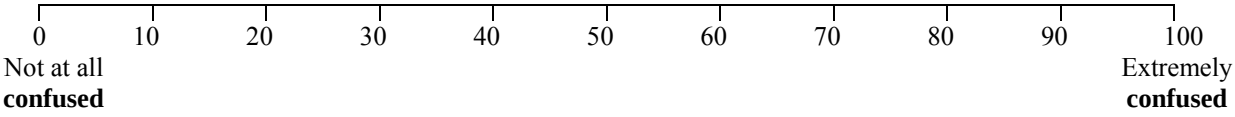
‘relaxed’



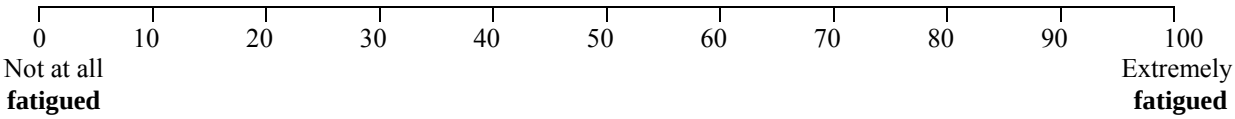
‘lively’



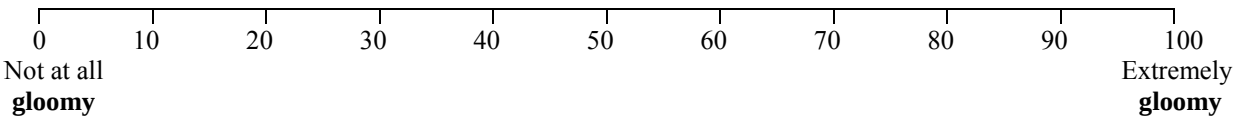
‘confused’



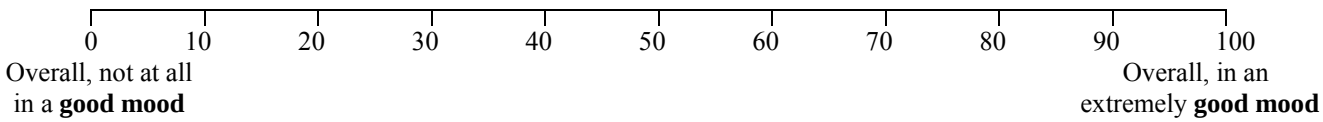
‘fatigued



‘gloomy’



‘overall mood’



[*In the version of the NRS used to assess general relaxation and mood over the past month, this wording is changed to ‘how you have felt in general over the past month’]

III. Associations between mood scales

The composite NRS for current mood, NRS-c, was created (1) to assess how comfortable and relaxed participants feel in this study, and also (2) as an attempt to simplify the POMS-SF and make it easier and quicker to complete, especially for those unfamiliar with some of the American English terms used in the POMS-SF (or the even the UK English words in the still shorter 24-item British Brunel Mood Scale, $BRUMS_{24}$ [Terry et al. 1999]). A project is currently underway to assess if the NRS can be validated against the $BRUMS_{24}$, which has in turn been validated against the [full version of] the POMS [Terry et al. 2003]. However, the NRS has not been validated against the POMS-SF directly, and that in part is the purpose of using both in this study.

In addition, the composite NRS for mood 'over the past month' , NRS-m, has been used as well as the Perceived Stress scales (PSS-10) to assess whether there is any association between them.

For the 17 participants who have so far used the questionnaires, the following correlations are significant (values are for Spearman's *rho*):

NRS item	**Anxious**	**Comf'able**	**Confused**	**Fatigued**	**Gloomy**	**Lively**	**Relaxed**	**Good m**
PSS	0.509**			0.403**	0.311**	-0.272**		-0.440**
POMS-1								
T	0.548**	-0.310+					-0.284+	-0.738**
D			0.298+	0.486**	0.724**	-0.535**		
A		-0.374*		0.532**	0.654**	-0.654**	-0.283+	-0.341+
V	-0.379*			-0.438*		0.564**	-.592**	
F		-0.349+	0.530**	0.911**	0.513**	-0.593**		-0.294+
C		-0.328+	0.546**	0.701**		-0.342+		-0.310+
Tot m d	0.365*	-0.234*	0.374*	0.805**	0.423*	-0.632**	-0.473**	-0.385*
POMS-2								
T	0.490**	-0.389*	0.399*	0.432*	0.582**		-0.355+	-0.459*
D	0.534**	-0.435*			0.686**	-0.448*		-0.296+
A	0.341+			0.331+	0.621**	-0.375*		
V		0.772**		-0.331+		0.714**	0.911**	0.475**
F	0.279+	-0.441*		0.885**		-0.395*	-0.457*	-0.640**
C	0.374*	-0.331+	0.405*	0.709**	0.422*		-0.347+	-0.645**
Tot m d	0.337+	-0.583**		0.759**	0.418*	-0.534**	-0.595**	-0.667**

** p<0.001; * p<0.01; + p<0.05

Key:

The POMS-SF subscales are:

T Tension-Anxiety
D Depression-Dejection
A Anger-Hostility
V Vigour-Activity
F Fatigue-Inertia
C Confusion-Bewilderment
Tot Total mood disturbance (a composite score based on the subscale scores)

Because in an earlier Pilot study the 'Anger-Hostility' dimension did not appear to be particularly affected by EA or TEA, it was omitted from the composite NRS in the interests of brevity.

The values in **bold** type are those where corresponding or opposed dimensions of the two scales intersect (e.g. 'Vigour-Activity' and 'Lively') and so would be expected to be high. This is the case for the two 'Confusion' and 'Fatigue' axes, for 'Depression' and 'Gloomy', for 'Anxiety' (the first time the NRS and POMS are used, but less so the second time), for 'Vigour' (the second time the NRS and POMS are used, but less so the first time), and for 'Total mood disturbance' (again the second time the NRS and POMS are used, but less so the first time). All the correlations in bold, except for Confusion (second use) and Total mood disturbance (first use) are significant at the 0.001 level. There are some other interesting cross-correlations between items belonging to different subscales.

It is reassuring that the PSS-10 correlates quite well with the 'Anxious' item in the NRS-m, but the lack of a negative correlation between the PSS-10 and the 'Relaxed' item invites further exploration.

Further study with more participants is underway to assess to what extent the composite NRS is valid when compared to the POMS-SF and the $BRUMS_{24}$ scale. Analysis of content validity, inter-rater reliability, split-half reliability, internal consistency reliability, and test-retest reliability (in an untreated group) are also envisaged, as described elsewhere [Mayor & Steffert 2013b].

IV. Methods used to create surrogate 'short slot' data

25% of data (19556 data points) was removed from AB4, a slot of the correct length, in 4 ways:

1. In 4 long blocks of unequal length, randomly distributed throughout the slot (this situation could arise if recording stalled and then restarted during data streaming, for example)

2. Evenly spaced throughout, in short strings < 1 sec (99 strings of 196 data pts, one of 152 data points, every 780 data pts)

3. In regularly spaced strings of increasing length from 0.1 sec to 10 sec.
28 multiples of 25 data points were removed (from 25 to 700), then 9 of increasing length, but not in multiples of 25, and a 38th consisting of 436 data points. They were removed every 2000 points.

4. In randomly spaced strings of random length from 2 sec to 0.5 sec.

Random numbers were used in Excel to generate 59 sequences of between 146 and 512 data points, in random order (totalling 19556 data points). These were allocated to randomly generated start points throughout the whole slot, from 231 points starting at 946 to 182 points starting at 77804. Where this resulted in overlapping removals, the more recent one was shifted to a start point at a multiple of 500 in one of the longer segments where no other data points had previously been removed (this occurred 8 times).