



OPEN

Neurophysiological and behavioural markers of compassion

Jeffrey J. Kim¹✉, Stacey L. Parker¹, James R. Doty², Ross Cunnington¹, Paul Gilbert^{1,3} & James N. Kirby¹

The scientific study of compassion is burgeoning, however the putative neurophysiological markers of programs which actively train distress tolerance, such as Compassionate Mind Training (CMT), are less well known. Herein we offer an integrative, multi-method approach which investigated CMT at neural, physiological, self-report, and behavioural levels. Specifically, this study first assessed participants' neural responses when confronted with disappointments (e.g., rejection, failure) using two fundamental self-regulatory styles, self-criticism and self-reassurance. Second, participant's heart-rate variability (HRV) – a marker of parasympathetic nervous system response – was assessed during compassion training, pre- and post- a two-week self-directed engagement period. We identified neural networks associated with threat are reduced when practicing compassion, and heightened when being self-critical. In addition, cultivating compassion was associated with increased parasympathetic response as measured by an increase in HRV, versus the resting-state. Critically, cultivating compassion was able to shift a subset of clinically-at risk participants to one of increased parasympathetic response. Further, those who began the trial with lower resting HRV also engaged more in the intervention, possibly as they derived more benefits, both self-report and physiologically, from engagement in compassion.

Access to mental health treatments have improved drastically over the last two decades, and yet rates of mental health problems have not decreased, with some research indicating they are on the rise^{1–3}. How does one make sense of this apparent paradox? Competing explanations suggest population growth, lack of treatment engagement, aging, and the possibility that although 12-month prevalence rates might not have decreased, length of mental health episodes have⁴. We offer an alternate possibility – namely the role of pathogenic self-criticism within modern society – which has increased the number of individuals with mental health issues, and also blunted the impact of available treatments on mental health problems. Importantly, the clinical literature is clear – there is substantial evidence that self-criticism contributes to mental health problems and impedes recovery^{5,6}.

In a recent major review of 48 studies, self-criticism is associated with: eating disorders, depressive disorders, social anxiety disorder, personality disorder, psychotic symptoms and interpersonal difficulties⁶. It is also associated with self-harm and suicide risk⁶. Self-criticism interferes with goal pursuit⁷, and is associated with problematic perfectionism, which has been increasing steadily over the last 20 years⁸. Ever since Freud argued that depression arose from 'anger turned inward'⁹, i.e., a form of self-criticism, it has been a therapeutic target for many schools of therapy¹⁰. Furthermore, Whelton and Greenberg¹¹ showed that it was when self-criticism generated hostile emotion to the self, that individuals felt particularly overwhelmed and defeated by their self-criticism. Hence, it's not just the content of a self-criticism that can be pathogenic, but the emotional intensity in which the criticism is delivered and experienced by the self that is important.

Exploring the forms and functions of self-criticism, Gilbert and colleagues¹² found two separate forms and functions relating to criticism. One form concerns a sense of inadequacy where the function is to improve oneself and avoid making mistakes, and the other form is a self-hating where the function is to punish or remove undesired aspects of the self. Furthermore, a study of 279 outpatients with mental health problems¹³ confirmed the importance of distinguishing self-hatred from self-improvement forms of self-criticism. These researchers also identified that self-hatred was a key mediator of the relationship between shame and poor mental-health outcomes¹³. While the forms and functions of self-criticism can be explored with self-report and clinical observation, research on the neurophysiological processes by which self-criticism exerts its pathogenic effects have also been

¹School of Psychology, The University of Queensland, Brisbane, Queensland, Australia. ²Center for Compassion and Altruism Research and Education, Stanford University School of Medicine, 2490 Hospital Drive, Suite 106, Stanford, California, USA. ³School of Allied Health and Social Care, University of Derby, Derby, DE22 1GB, United Kingdom. ✉e-mail: Jeffrey.kim@uqconnect.edu.au

explored. In an fMRI study, Doerig *et al.*¹⁴ invited participants to focus on self-critical versus neutral adjectives. They found activation of dorsomedial and lateral frontal pathways, as well as the amygdala and posterior cingulate, when engaged in self-criticism. Longe *et al.*¹⁵ used a different methodology and presented participants with written stimuli which describe a mistake, setback, or failure, and manipulated the emotionality of the vignettes (i.e., negative versus neutral) as well as the perspective to adopt (i.e., self-criticism versus self-reassurance). These factors were associated with unique neural substrates, whereby self-criticism was associated with activation within the DLPFC, and self-reassurance the anterior insula¹⁵.

Whilst important research into the neural and self-report bases of criticism have been conducted, however, research has yet to continue an understanding of factors which may dampen criticism's pathogenic effects, such as self-compassion and/or self-reassurance. Self-reassurance is the ability to be encouraging, resilient, and aware of one's positive qualities and capabilities in the face of setbacks, mistakes or failures¹⁶. Importantly, the ability for self-reassurance has been shown to protect against symptoms of psychopathology, such as depression and anxiety¹⁶, whereas self-criticism has been shown to predispose toward their development¹⁷. Within a Compassionate Mind Training (CMT) framework, self-reassurance is the cognitive relating style used when adopting a compassionate-mind¹⁶. Compassion is defined as, "a sensitivity to suffering in self and others, with a commitment to try and alleviate and prevent it (p. 2)¹⁸". Thus, if the motive of a behavior or cognitive relating style is to reduce suffering there are many ways one can do this, such as being kind to oneself, as well as being self-reassuring^{16,18}.

Within the present research, we will replicate and extend Longe and colleague's fMRI paradigm¹⁵ by incorporating a focal emotion contrast (i.e., negative minus neutral), to further examine how self-criticism versus self-reassurance might modulate neural responses to negative emotion. In addition, our integrative and multi-method approach, which first assessed participants' neural markers of self-critical versus self-reassuring motivations with fMRI, also tracked the same participants through a compassionate mind training protocol with markers of physiology, specifically heart-rate variability (HRV), assessed pre- and post- a two-week self-directed engagement period.

Extensive neuroimaging research has associated activation of a core salience network (SN), involved in the processing of affective, painful stimuli^{19–21}. Comprised of two key regions, the anterior insula and anterior cingulate, activation of this network has been implicated across pre-existing compassion and empathy fMRI studies to date²². Importantly, whilst affective stimuli have been shown to produce activation in this network, cultivating compassion has been shown to reduce the salience and experience of negative emotions²³. Therefore, we expect to observe activation of this network for emotional written vignettes and anticipate that self-criticism would boost this activation whereas self-reassurance would suppress this activation. An additional network reputed to be involved in compassion-related processes is the default-mode network (DMN)^{24,25}, jointly involved in self-referential and mentalizing processes^{25,26}. Critically, recent efforts have sought to understand the role of default-mode function in clinical interventions²⁷, specifically with forms of mental training, such as meditation, which have been associated with alterations in DMN activity and functional connectivity²⁸. Therefore, we anticipate potential recruitment of this network in parallel with SN activation when engaged in self-criticism and self-reassurance motivations, given our task is self-referential and mentalizing in nature.

Next, in our brief Compassionate Mind Training (CMT) intervention, we utilized the practice "Cultivating the Compassionate Self", which is a central compassionate mind training practice within Compassion Focused Therapy (CFT). This intervention comprises four core elements; (1) grounding and body posture, (2) soothing-rhythm breathing, (3) mindfully noting when self-criticism arises, and switching to (4) cultivating the compassionate-self²⁹. With this training protocol, we assessed participant's HRV at both rest and during the compassion exercise. Higher HRV is indicative of higher parasympathetic nervous system outflow via the Vagus nerve^{30–33}, activity which have been associated with increased states of contentment, calmness, safety³¹. Notably, when the sympathetic nervous system is recruited for defensive or other behavioural activation systems, HRV is reduced³¹. Previous research has identified how these CMT exercises can significantly increase baseline HRV and reduce symptoms of depression, anxiety, stress, and fears of self-compassion^{34,35}. Hence, we predict our compassion intervention will, compared to baseline, increase HRV. As an additional exploratory analysis, we will examine the percentage improvement of participants who move out of a clinical 'at risk' range of HRV when engaged in compassion, as identified in a recent paper which assigned elevated cardiac risk to low resting HRV³⁶. We will also assess the degree to which two-weeks of training is sufficient to reduce fears of self-compassion, and we will explore the relationship between HRV and behavioural markers of engagement in compassion training across the two-week self-directed engagement period, as assessed via frequency of accessing the training link.

Results

Neural activity for imagery toward emotional stimuli. To test whether emotionally salient stimuli (written vignettes) elicited activation in brain regions which code for negative affect, we examined brain activity during imagery toward emotional minus neutral statements. Within-subjects contrasts revealed significantly greater activation for emotional compared with neutral stimuli, tested across the whole-brain, found in regions of the SN and DMN, as well as regions of the occipital cortex (see Fig. 1a, all $pFWEs < 0.05$). This neural activation included: left PCC, left calcarine gyrus, left and right lingual gyrus, left MPFC, and left ACC (See Supplementary Table 1 for peak co-ordinates and statistical values).

Neural activity for imagery during self-criticism and self-reassurance. Next, we conducted an ROI approach to examine differences in neural activation when participants engaged in self-criticism and self-reassurance. We extracted neural responses from the ACC, MPFC, PCC and the lingual gyrus, to extend upon significant whole-brain activation as identified above. We also extracted an ROI from the left amygdala and left insula cortex, as previous research across multiple studies have linked activity in these regions with

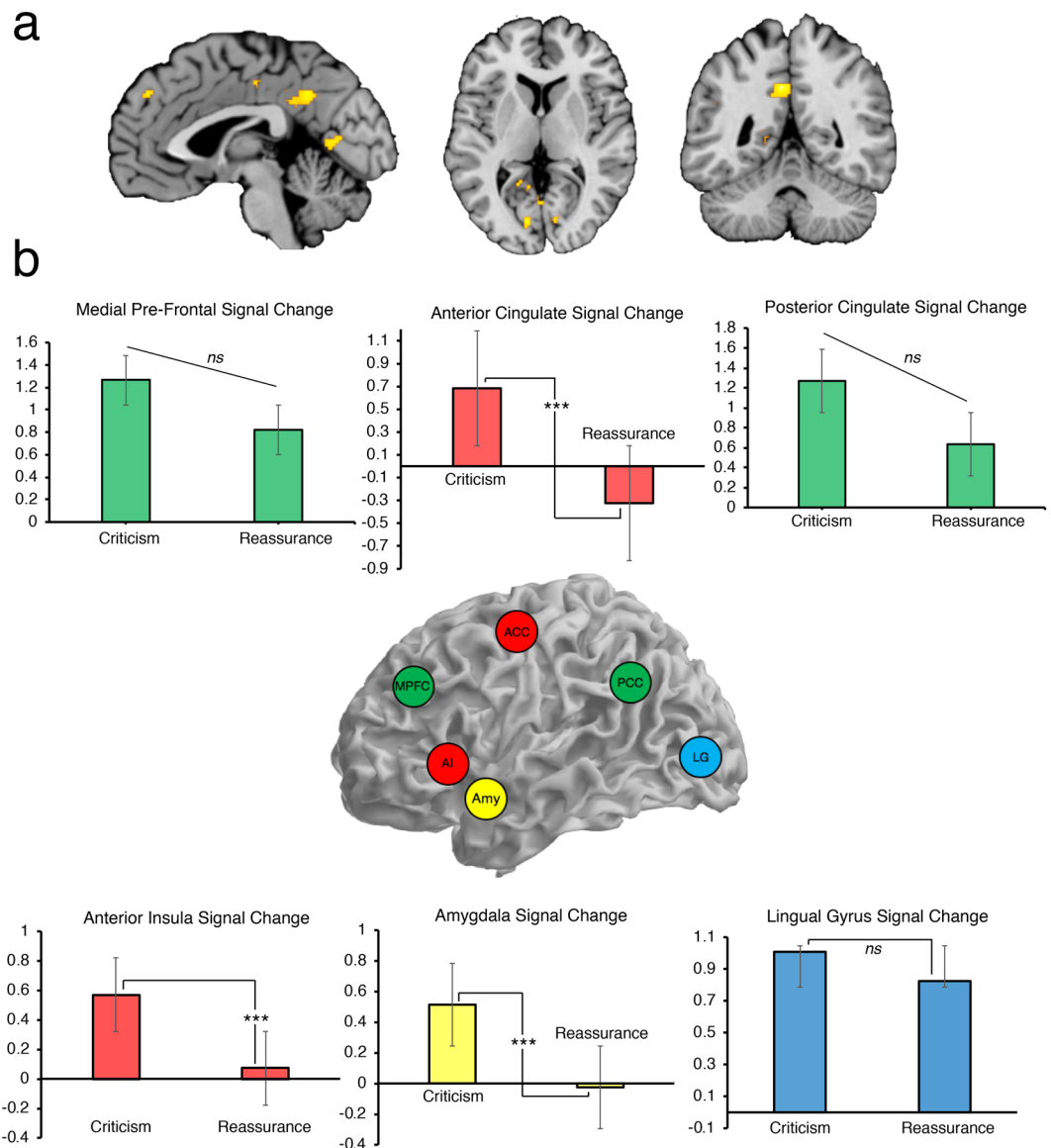


Figure 1. Thresholded whole-brain activation for the Emotional – Neutral contrast, revealing an overall effect of emotional imagery, and ROI approach comparing criticism to reassurance. **(a)** Left: Sagittal image of Posterior Cingulate, Lingual Gyrus, and Medial Frontal Cortex ($x = -2$). Middle: Axial image of Lingual and Calcarine Gyri ($z = 4$). Right: Coronal image of Posterior Cingulate ($y = -50$). Activation corrected at $p_{FWE} < 0.05$, and coordinates reported in MNI-space. **(b)** ROI approach comparing self-critical imagery to self-reassuring imagery, with % signal change on the Y-axis. Top: Comparison of MPFC (2 46 36), ACC (0 14 36), and PCC (4 52 36) % signal change. Middle: Inset which depicts rough spatial location of ROIs. Bottom: Comparison of Left AI ($-26\ 10\ -14$), Left Amygdala ($-28\ -4\ -12$), and Middle Lingual Gyrus (0 $-68\ 6$) % signal change. Coordinates reported in MNI-space. Error bars indicate standard error. $N = 40$. *** $p < 0.001$, $ns\ p > 0.05$.

compassion-related processes^{22,37,38}. Voxel-wise coordinates for each ROI were defined as each participant's peak cluster identified with the AAL labelling tool under the WFU-pick atlas in SPM12, from the emotional-neutral contrast, for both self-criticism and self-reassurance. As can be seen in Fig. 1b, significantly greater neural activation within the AI, ACC, and the amygdala, was identified for self-criticism as compared with self-reassurance (paired-sample t-tests, all p 's < 0.001). Activity within the MPFC, PCC, and lingual gyrus did not differ between conditions (Fig. 1b, paired sample t-tests, $p > 0.05$, ns).

Physiology at rest versus intervention. To derive a measure of parasympathetic response with our HRV, we created natural-logarithm units of high-frequency components (0.15–0.40 Hz) of participants ECG-trace, for each 15-minute rest and 15-minute intervention component, at both pre- and post- two-week training. As can be seen in Fig. 2, a two-way ANOVA revealed a significant within-subjects effect of session (i.e., rest vs intervention), whereby HRV was higher when engaged in the compassion intervention than at rest ($F(1,36) = 5.673$,

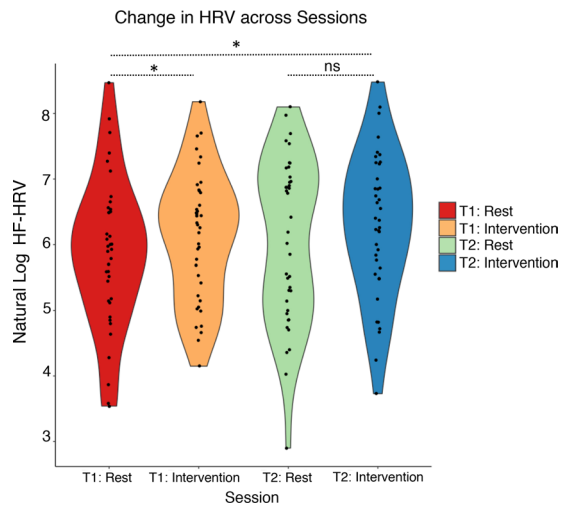


Figure 2. Overall changes in physiology (HRV) derived from our compassion intervention. Changes in HRV, expressed as natural log units, are depicted at rest and during intervention, at both pre- and post- two-week intervention. $N = 38$, accounting for listwise deletion. $*p < 0.05$, $ns p > 0.05$.

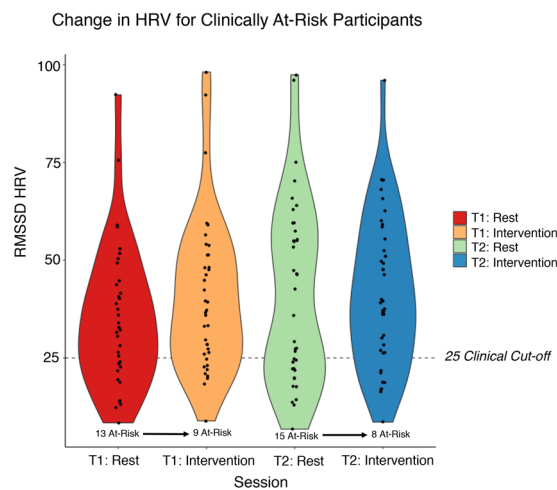


Figure 3. Percentage improvement of a subset of participants shifted out of a clinically at-risk range of poor resting-HRV, after engagement in compassion. X-axis depicts pre- and post- two-week training, at either rest or intervention (compassion meditation), and Y-axis depicts RMSSD HRV value. Dashed-line depicts clinical cut-off (25), and subplot text describes the number of clinically at-risk participants which moved above the clinical cut-off after engagement in compassion. $N = 38$, accounting for listwise deletion.

$p < 0.023$). However, we observed a non-significant within subjects effect of order (i.e., pre and post two-week training), $F(1,36) = 2.694$, $p = 0.109$, whereby rest versus intervention did not differ between timepoints, and a non-significant interaction ($F(1,36) = 0.033$, $p = 0.857$). Follow-up t-tests revealed significant increases in HRV from T1 rest to T1 intervention ($t(37) = 2.310$, $p < 0.027$), as well as from T1 rest to T2 intervention ($t(36) = 2.667$, $p < 0.011$). T2 rest to T2 intervention was non-significant ($t(38) = 1.813$, $p = 0.078$).

HRV-shift for clinically at-risk participants. Next, we assessed the percentage improvement of participants who were able to shift out of a clinically at-risk range of low resting-HRV, as defined empirically³⁶. In accordance with these guidelines, we report raw RMSSD HRV scores for interpretation, and note RMSSD HRV is considered a reliable and valid time-series metric of vagally-mediated HRV^{33,39}. As can be seen in Fig. 3, four and seven clinically-at risk participants were shifted above the clinical cut-off of 25 RMSSD³⁶ when engaged in the compassion intervention versus at rest, at either pre- and post- two-week training, respectively.

Relationship between physiology and behaviour. Next, we examined the relationship between HRV and how often participants engaged in compassion training across the two-weeks. From behavioural metrics of how often participants accessed the compassion exercise across the two-week period, which we assessed as frequency of recording access from the experiment website, we median split the data into those who engaged More

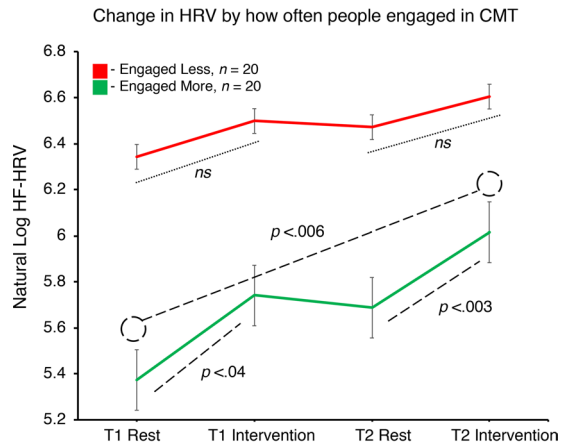


Figure 4. Relationship of HRV and Behaviour (frequency of engagement in compassion), whereby data were median split into those with Less versus More engagement of the compassion recording across the two-week engagement period. $N = 38$, accounting for listwise deletion. Error bars indicate standard error. $**p < 0.01$, $*p < 0.05$, $ns p > 0.05$.

versus Less. Those who engaged Less ($N = 20$) accessed the recording on average 3 times across the two-week period ($M = 2.9$, $SD = 1.2$, range = 1–4), in contrast, those who engaged More ($N = 20$) accessed the recording on average 9 times across the two-week period ($M = 9.1$, $SD = 3.1$, range = 5–15). A 3-way mixed ANOVA revealed a significant between-subjects effect, whereby those who engaged Less had higher HRV overall, $F(1,35) = 6.746$, $p < 0.014$ (Fig. 4). We observed a non-significant effect of session order, whereby HRV pre- versus post- were not significantly different ($p > 0.05$, ns). The three-way interaction term between participant engagement, session type, and session time was non-significant ($p > 0.05$, ns). Follow-up paired t-tests revealed significant increases in HRV for those who engaged More across the two-weeks, from rest to intervention at Time 1 ($p < 0.04$), rest to intervention at Time 2 ($p < 0.003$), and from Time 1 rest to Time 2 intervention ($p < 0.006$). These effects were non-significant for those who engaged Less (p 's > 0.05 , ns).

Pre-post self-report differences. Finally, we observed a significant decrease in fears of expressing compassion to the self⁴⁰. Paired sample t-test (Time 1: $M = 54.53$, $SD = 22.83$, Time 2: $M = 33.68$, $SD = 15.69$), $t(39) = 5.683$, $p < 0.001$.

Discussion

This study conducted an integrative, multi-method approach which first investigated two fundamental self-regulatory styles (self-criticism and self-reassurance) with fMRI, and second, measured participant's HRV, a marker of parasympathetic response, during compassion training, pre- and post- a two-week self-directed training period.

Upon investigation of neural correlates of self-critical versus self-reassuring motivations, our manipulation check for a focal emotion contrast identified increased neural activation across SN and DMN, as well as foci within the visual cortex (Fig. 1a). Importantly, a ROI approach revealed self-reassurance and self-criticism modulated neural responses to negative emotion, in regions such as the AI, ACC, and amygdala (Fig. 1b), whereby self-reassurance was associated with a decrease in neural reactivity within these regions, as opposed to self-criticism, which boosted neural activation within these regions. Interestingly, activity in MPFC, PCC, and lingual gyrus (within visual cortex) did not differ between self-criticism and self-reassurance.

Overall, these findings replicate previous research which identified SN activity as associated with compassion-related processes^{19,22}, and we are the first to explicitly afford the DMN a putative role for compassion-related processes, both self-critical and self-reassurance motivations. Interestingly, in direct contrast with the results of Longe *et al.*¹⁵, our data show that both criticism and reassurance modulate similar neural pathways, rather than operating on distinct neural regions (i.e., DLPFC) or networks (i.e., SN). However, we note the focus of the current experiment were to identify neurophysiological correlates of compassion and criticism, as well as physiological shift during compassion cultivation. More work is needed to directly test for similarities and/or differences in these neural systems, as this was not the aim of the current experiment.

Second, with our CMT protocol, we identified increased HRV during compassion training versus at rest, indicative of increased parasympathetic response (Fig. 2). Overall these findings replicate previous research which has identified increased HRV during meditation practices^{34,35,41}, such as compassion and mindfulness training. In addition, our brief compassionate mind training protocol was also able to shift a subset of participants out of a clinically at-risk range of 25 RMSSD HRV at rest (Fig. 3).

Third, we identified resting HRV predicted behavioral markers of engagement in compassion training across the two-weeks (i.e., levels of recordings accessed). Specifically, individuals who had lower resting HRV engaged more in the intervention and derived more physiological benefit (i.e., increases) in HRV from rest to intervention, at both pre- and post- two-week intervention (Fig. 4). In addition, we note also that individuals with higher resting HRV engaged less in compassion during the two-week training period, and did not derive significant

physiological shift from rest to intervention across pre- and post-training (Fig. 4). Last, our data also identified a significant reduction in fears of self-compassion after the two-week intervention, which provides a validation of the effects of CMT at the self-report level.

Importantly, our HRV data suggests individual differences in compassionate responding exist, whereby for some individuals significant HRV-change across sessions did not occur (Fig. 4). Participants never received feedback on their HRV scores or what this data may indicate, yet those with lower resting HRV were also those which engaged in the practice more frequently and benefitted physiologically as a result. Thus, this finding can potentially position our results in the context that individuals whom really needed the practice, physiologically, did use it. These findings may extend Gilbert's evolutionary and neurophysiologically-informed model of compassion²⁹, to highlight how individual differences across threat- versus self-soothing systems may relate to HRV-response and engagement. For example, CMT was designed as an adjunct to traditional psychotherapies for those whom are treatment non-responsive³⁴. The clinical implications of these data are fascinating - given our data show increases in HRV during a compassion meditation state, to what extent are these differences not apparent for those who might have fears, blocks or resistances to compassion⁴⁰? Furthermore, how intensive might training need to be in order to provide benefit to these individuals?

In accordance with our data that compassion and criticism may recruit similar neural networks and regions, we suggest that the act of imagery generated while engaging in self-criticism may act upon similar subcortical routes for processing threatening stimuli in general, processes which may be stimulated automatically and unconsciously⁴²⁻⁴⁴. Interestingly, previous research has identified subconscious pathways for threat processing in the amygdala and visual cortex (lingual gyrus) apparent even under conditions of cortical blindness^{45,46}, therefore we suggest that amygdala and lingual gyrus activation likewise identified within our research may subsume a similar route.

The aim of this integrative project was to isolate specific components of CMT, and to balance a multi-modal design with a brief two-week training component. Therefore, we utilized a within-groups approach as our *a priori* aim was to link neurophysiological, behavioural, and self-report data. However, to extend upon this work with a between groups design would be of benefit, to determine whether CMT is more effective than other mental training, such as mindfulness. We would argue that CMT training for threat would be more robust than mindfulness, given CMT is at its core comprised of training to try and be helpful to observed suffering and/or distress. Whereas mindfulness meditation may instead direct attention toward a mantra and/or alternate locus of attention⁴⁷, and to effectively dissociate from threat, given the nature of its attention training paradigm. We anticipate future work which could investigate physiological (HRV) shift to both safety and threatening scenarios would be a direct test of this hypothesis, with procedures such as the Trier Social Stress Test⁴⁸. Here, we anticipate that CMT would enable participants to better regulate threat and return to parasympathetic response more quickly after threat, as opposed to mindfulness training. Given recent research has identified that acting altruistically in the face of threat can reduce both physical and psychological pain⁴⁹, it is possible that CMT may likewise hold an analgesic or moderating effect on threat and pain.

Furthermore, although the median split approach has been questioned⁵⁰, we affirm the robustness of this technique given the presence of uncorrelated or non-linear related variables^{51,52}. Given we now better understand neurophysiological markers of compassion within a healthy control sample, however, we look forward to future research which could potentially consider a control versus a clinically-relevant sample (i.e., depression and anxiety). Our HRV data may tentatively support a distinction between these groups, first given participants' baseline HRV varies considerably, and second that we were able to shift a subset of at-risk participants above the clinical cut-off when engaged in soothing compassion states (Fig. 3). In addition, whilst we have identified the efficacy of CMT to down-regulate neural markers of threat and increase physiological and self-report makers of safety, soothing, and regulation, we note however that we did not observe a baseline HRV shift for participants at pre- to post-. Therefore, in order to better understand who may benefit from cultivating compassion, more work is needed, potentially with a more longform approach, to assess for a true baseline shift, in both "normal" and clinically relevant samples.

Methods

Participants. 40 participants (Mean age = 22 years, SD = 0.49, 27 female) took part in the present study. The University of Queensland Health and Behavioural Sciences, Low & Negligible Risk Ethics Sub-Committee approved the experimental protocol, and this project complies with the provisions contained in the *National Statement of Ethical Conduct in Human Research* and complies with the regulations governing experimentation on humans. Participation was voluntary and anonymous, and subjects provided informed, written and/or electronic consent. Experimental stimuli and procedures, analysis code, data, and statistical output are available in an OSF repository (10.17605/OSF.IO/6H9MD), and fMRI data are available on a University of Queensland Research Data Manager (RDM) server, with access available upon request.

fMRI Stimuli. We created 60 written stimuli in total, consisting of a personal mistake, setback or failure. 30 statements were of emotional valence whereas 30 were neutral (i.e., "I fail to keep up with my commitments in life", and "I keep up with my commitments in life", respectively). Our neutral stimuli were created to describe a non-emotive, non-intense control to counterbalance the emotional stimuli set. For both emotional and neutral sets we assessed two metrics, valence (1-5, where 1 = Very Unpleasant) and intensity (1-5, where 1 = Not Intense). Our emotional statements ($n = 30$) were revealed to be sufficiently unpleasant ($M = 1.89$) and intense ($M = 3.54$), with all neutral statements ($n = 30$) described as less unpleasant ($M = 3.80$) and comparatively not intense ($M = 2.34$).

fMRI design. Within the scanner pre- two-week training, we examined participant's neural responses to the validated (affective and neutral) written stimuli when engaged in self-criticism and self-reassurance. After each trial within a block of either self-criticism or self-reassurance, participants rated how intense their degree of self-criticism or self-reassurance was to each statement (button-press on an MR-compatible button box which ranged from 1–4, where 1 = not very intense, and 4, very intense). A typical trial consisted of stimuli presented for a 6 second duration, followed by a rating of intensity for a 3 second duration, and an inter-trial-interval of 0.5 seconds. The first order of instruction for a particular block, that is, self-reassurance verses self-criticism, was counterbalanced for a total of 8 blocks. As our focal contrast, we manipulated the emotionality of the statements within scan runs (“emotive” vs “neutral”), in a counterbalanced order across participants. 30 statements were quasi-randomized across participants and presented for a total of 30 trials per fMRI run (~6.5 min total duration) over a total of 8 repeated fMRI runs. Participants were given 10 practice trials of emotional and neutral stimuli, and rated stimuli on intensity.

HRV design. Approximately a week after the fMRI experiment, participants listened to a standardized 15-minute recording of the CFT exercise “Cultivating the Compassionate Self”, with HRV collected pre- and post- a two-week self-directed engagement period. The guided exercise was recorded in English by a Clinical Psychologist and expert in CFT. A website to facilitate this project during the two-week engagement period between Time 1 and Time 2 was created to house the audio recording and various instructions/helpful tip sheets for participants.

fMRI acquisition and pre-processing. We collected our fMRI data on a 3-Tesla Siemens Trio MRI scanner utilizing a 64-channel head-coil. A gradient-echo, echo-planar “fast imaging” (EPI) sequence were used to acquire functional images, with the following sequence parameters: 60 horizontal slices (2×2 -mm in-plane voxel resolution and 2-mm slice thickness plus 10% gap), repetition time (TR) 1000 ms; echo time (TE) 30 ms. Eight identical fMRI runs of 292 images (6 minutes each) were acquired. A 3D high-resolution, unified and denoised T1-weighted MP2RAGE image across the entire brain was also acquired and used as anatomical reference for subsequent pre-processing in SPM12 (TR = 4000 ms, TE = 2.93 ms, FA = 6° , 176 cube matrix, voxel size = 1-mm). Functional imaging data were pre-processed and analyzed using SPM12, implemented in MATLAB. Structural T1-scans were co-registered to the average of the spatially realigned functional slices. Next, an inbuilt segmentation routine was applied to register each structural T1-image to the standard MNI template in MNI space. These transform parameters elicited from segmentation were subsequently applied to all realigned images, resliced to a $2 \times 2 \times 2$ -mm resolution and smoothed with 6-mm full-width-at-half-maximum (FWHM) isotropic Gaussian kernel.

fMRI first and second-level analyses. For first-level data analysis, block-related neural responses to stimuli were modelled as 2 separate conditions (all combinations of emotional/neutral, self-criticism/self-reassurance) and convolved with the canonical hemodynamic response function (HRF). For group level analysis, wholebrain contrasts of emotional-neutral stimuli overall were reported at a cluster-level threshold of $p < 0.05$, corrected for family-wise error, with clusters formed with a voxel-level height threshold at $p < 0.001$, uncorrected. Brain regions shown to be significant had their anatomical labels identified with the Automated Anatomical Labelling (AAL) toolbox implemented in SPM12⁵³. Next, in order to examine correlations between the level of neural activation (i.e. difference in response between emotion verses neutral) and the mindset participants engaged in (i.e. self-criticism versus self-reassurance), we performed additional region of interest (ROI) analyses. For each ROI, we identified peak clusters which showed significantly greater activation overall for emotion vs neutral stimuli, and used these coordinates to extract the average contrast parameter estimates (i.e. levels of activation, Beta weights) with 5-mm radius spheres centered on those peaks for each mindset (i.e., self-criticism and self-reassurance). We then used SPSS to examine the correlation between neural responses and the mindset participants engaged in when processing neural responses to emotional stimuli, and used the R Statistical Programming Language to visualize both regions-of-interest results, as well as subsequent HRV data.

fMRI trial-by-trial intensity ratings. Analysis of participant's mean level of intensity for reassurance (emotional statements: $M = 2.45$, $SD = 0.48$, neutral statements: $M = 2.63$, $SD = 0.64$) and criticism (emotional statements: $M = 2.92$, $SD = 0.45$; neutral statements: $M = 2.07$, $SD = 0.52$) revealed intensity ratings were significantly higher for critical (emotional – neutral) but not for reassuring (emotional – neutral) trials ($t(38) = 7.300$, $p < 0.001$, and $t(38) = -1.372$, $p = 0.178$, *ns*, respectively). Inspection of a correlation matrix revealed intensity ratings for reassurance and criticism during emotional trials were correlated ($R = 0.47$, $p < 0.003$).

HRV data acquisition, pre-processing and analysis. We collected our heart-rate variability (HRV) data with a portable Schiller Systems Medilog electrocardiogram (ECG) device. These recorders of the “AR12plus” type utilize a 3-lead bipolar ECG amplifier, and we positioned each lead on the chest, to maximize ECG amplitude and minimize muscle artifacts. After data collection, data were digitized, pre-processed, and extracted using Schiller's proprietary Medilog Darwin software package, which processes the ECG biosignal to comply with accepted guidelines in the field for measurement and analysis of HRV data. Via visual inspection of the ECG biosignal, and after application of an algorithm which automatically filters for normal and non-normal beats, we excluded signal artifacts and irregular (non-normal) beats not removed from the initial application of the pre-processing algorithm. Therefore only the normal beats (via R-wave peak detection) were extracted. The sampling period was 1000 Hz at a time resolution of 250 μ s (microseconds), and we utilized linear detrending. After initial pre-processing, high-frequency (HF) HRV components (i.e., 0.15–0.40 Hz) of each participants' ECG trace were extracted upon application of the discrete Fourier transform (DFT) of the beat-to-beat interval times-series. We also extracted raw RMSSD HRV components, namely the Root Mean Square of Successive Differences of

the HRV beat-to-beat timeseries. Data were coded as being from either Time 1 or Time 2, at either Rest or Intervention. Signal averages for each of these periods were created from the bins downloaded from Medilog Darwin (i.e., 4*15-minute bins, across Rest and Intervention, at both Time 1 and Time 2). We did not measure or control for respiration during the ECG recording, yet we did attempt to control for additional sources of variance such as posture, exercise, alcohol, caffeine intake, and time of testing. HRV data were transformed to the natural logarithm (base of the mathematical constant e) in order to better approximate normality. Missing data for HF-HRV and RMSSD were less than 0.03% and 0.04%, overall, and missing data were counted as listwise deletion in all relevant statistical tests.

Received: 13 February 2020; Accepted: 6 April 2020;

Published online: 22 April 2020

References

1. Furukawa, T. A. & Kessler, R. C. Why has prevalence of mental disorders not decreased as treatment has increased? *Aust. N. Z. J. Psychiatry* **53**, 1143–1144 (2019).
2. World Health Organization. World health report: Mental disorders affect one in four people, http://www.who.int/whr/2001/media_centre/press_release/en/, Published 2001. Accessed January 28, (2020).
3. World Health Organization. Mental Health ATLAS 2017. Member State Profile: United States, https://www.who.int/mental_health/evidence/atlas/mental_health_atlas_2017/en/, Published 2017. Accessed January 28, (2020).
4. Harvey, S. B. *et al.* Is the prevalence of mental illness increasing in Australia? Evidence from national health surveys and administrative data. *Med. J. Aust.* **206**, 490–493 (2017).
5. Rose, A. V. & Rimes, K. A. Self-criticism self-report measures: Systematic review. *Psychol. Psychother. Theory, Res. Pract.* **91**, 450–489 (2019).
6. Werner, A. M., Tibubos, A. N., Rohrmann, S. & Reiss, N. The clinical trait self-criticism and its relation to psychopathology: A systematic review. *J. Affect. Disord.* **246**, 530–547 (2019).
7. Powers, T. A., Koestner, R., Zuroff, D. C., Milyavskaya, M. & Gorin, A. A. The effects of self-criticism and self-oriented perfectionism on goal pursuit. *Personal. Soc. Psychol. Bull.* **37**, 964–975, <https://doi.org/10.1177/0146167211410246> (2011).
8. Curran, T. & Hill, A. P. Perfectionism is increasing over time: A meta-analysis of birth cohort differences from 1989 to 2016. *Psychol. Bull.* **145**, 410–429, <https://doi.org/10.1037/bul0000138> (2019).
9. Freud, S. Trauer und Melancholie. *Internationale Z. für Ärztliche Psychoanalyse* **4**, 288–301 (1917).
10. Kannan, D. & Levitt, H. M. A review of client self-criticism in psychotherapy. *J. Psychother. Integr.* **23**, 166–178, <https://doi.org/10.1037/a0032355> (2013).
11. Whelton, W. J. & Greenberg, L. S. Emotion in self-criticism. *Pers. Individ. Dif.* **38**, 1583–1595, <https://doi.org/10.1016/j.paid.2004.09.024> (2005).
12. Gilbert, P., Clarke, M., Hempel, S., Miles, J. N. V. & Irons C. Criticizing and reassuring oneself: An exploration of forms, styles and reasons in female students. **43**, 31–50, <https://doi.org/10.1348/01446650472812959> (2004).
13. Castilho, P., Pinto-Gouveia, J. & Duarte, J. Two forms of self-criticism mediate differently the shame–psychopathological symptoms link. *Psychol. Psychother. Theory, Res. Pract.* **90**, 44–54, <https://doi.org/10.1111/papt.12094> (2017).
14. Doerig, N. *et al.* Neural representation and clinically relevant moderators of individualised self-criticism in healthy subjects. *Soc. Cogn. Affect. Neurosci.* **9**, 1333–1340, <https://doi.org/10.1093/scan/nst123> (2013).
15. Longe, O. *et al.* Having a word with yourself: Neural correlates of self-criticism and self-reassurance. *Neuroimage*. **2010** **49**, 1849–1856, <https://doi.org/10.1016/j.neuroimage.2009.09.019> (2010).
16. Petrocchi, N., Dentale, F. & Gilbert, P. Self-reassurance, not self-esteem, serves as a buffer between self-criticism and depressive symptoms. *Psychol. Psychother. Theory, Res. Pract.* **92**, 394–406, <https://doi.org/10.1111/papt.12186> (2019).
17. Gilbert, P. & Procter, S. Self-Criticism: Overview and Pilot Study of a Group Therapy Approach. *Clin. Psychol.* **379**, 353–379, <https://doi.org/10.1002/cpp> (2006).
18. Gilbert, P. *et al.* The development of compassionate engagement and action scales for self and others. *J. Compassionate Heal. Care* **4**, 1–24, <https://doi.org/10.1186/s40639-017-0033-3> (2017).
19. Hein, G., Engelmann, J. B., Vollberg, M. C. & Tobler, P. N. How learning shapes the empathic brain. *Proc. Natl Acad. Sci. U S A* **113**, 80–85 (2016).
20. Uddin, L. Q., Yeo, B. T. T. & Spreng, R. N. Towards a Universal Taxonomy of Macro-scale Functional Human Brain Networks. *Brain Topogr.* **1–17**, <https://doi.org/10.1007/s10548-019-00744-6> (2019)
21. Rutgen, M. *et al.* Placebo analgesia and its opioidergic regulation suggest that empathy for pain is grounded in self pain. *Proc. Natl Acad. Sci. U S A* **112**, 638–646 (2015).
22. Kim, J. J., Cunnington, R. & Kirby, J. N. The neurophysiological basis of compassion: An fMRI meta-analysis of compassion and its related neural processes. *Neurosci. Biobehav. Rev.* **108**, 112–123 (2020).
23. Arimitsu, K., & Hofmann, S. G. Effects of compassionate thinking on negative emotions. *Cognition and Emotion*, **31**, 160–167 <https://doi.org/10.1080/02699931.2015.1078292> (2017).
24. Raichle, M. E. *et al.* A default mode of brain function. *Proc. Natl Acad. Sci. U S A* **98**, 676–682 (2001).
25. Gusnard, D. A., Akbudak, E., Shulman, G. L. & Raichle, M. E. Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proc. Natl Acad. Sci. U S A* **98**, 4259–4264 (2001).
26. Amodio, D. M. & Frith, C. D. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* **7**, 268–277 (2006).
27. Andrews-Hanna, J. R., Smallwood, J. & Spreng, R. N. The default network and self-generated thought: component processes, dynamic control, and clinical relevance. *Ann. N. Y. Acad. Sci.* **1316**, 29–52 (2014).
28. Brewer, J. A. *et al.* Meditation experience is associated with differences in default mode network activity and connectivity. *Proc. Natl Acad. Sci. U S A* **108**, 20254–20259 (2011).
29. Gilbert, P. The origins and nature of compassion focused therapy. *Br. J. Clin. Psychol.* **53**, 6–41 (2014).
30. Porges, S. W. The Polyvagal Perspective. *Biol. Psychol.* **74**, 116–143 (2007).
31. Kirby, J. N., Doty, J. R., Petrocchi, N. & Gilbert, P. The current and Future role of Heart rate variability for Assessing and training compassion. *Front. public. Heal.* **5**, 1–6 (2017).
32. Task Force. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Eur. Heart J.* **17**, 354–381 (1996).
33. Laborde, S., Mosley, E. & Thayer, J. F. Heart rate variability and cardiac vagal tone in psychophysiological research - Recommendations for experiment planning, data analysis, and data reporting. *Front. Psychol.* **8**, 1–18 (2017).
34. Matos, M. *et al.* Psychological and Physiological Effects of Compassionate Mind Training: a Pilot Randomised Controlled Study. *Mindfulness* **8**, 1699–1712 (2017).

35. Petrocchi, N., Ottaviani, C. & Couyoumdjian, A. Compassion at the mirror: Exposure to a mirror increases the efficacy of a self-compassion manipulation in enhancing soothing positive affect and heart rate variability. *J Posit Psychol.* **12** (2017).
36. Jarczok, M. N., Koenig, J., Wittling, A., Fischer, J. E. & Thayer, J. F. First Evaluation of an Index of Low Vagally-Mediated Heart Rate Variability as a Marker of Health Risks in Human Adults: Proof of Concept. *J. Clin. Med.* **8**, 1–12 (2019).
37. Kedia, S., Berthoz, M., Wessa, D. & Hilton, J. Martinot. An agent harms a victim: a functional magnetic resonance imaging study on specific moral emotions. *J. Cogn. Neurosci.* **20**, 1788–1798 (2008).
38. Mascaro, J. S., Rilling, J. K., Tenzin, N. L. & Raison, C. L. Compassion meditation enhances empathic accuracy and related neural activity. *Soc. Cogn. Affect. Neurosci.* **8**, 48–55 (2014).
39. Moss, D. Heart rate variability (HRV) Biofeedback. *Psychophysiol. Today* **1**, 1–336 (2004).
40. Kirby, J. N., Day, J. & Sagar, V. The ‘Flow’ of compassion: A meta-analysis of the fears of compassion scales and psychological functioning. *Clin. Psychol. Rev.* **70**, 26–39 (2019).
41. Ceccarelli, L. A., Giuliano, R. J., Glazebrook, C. M. & Strachan, S. M. Self-Compassion and Psycho-Physiological Recovery From Recalled Sport Failure. *Front. Psychol.* **10**, 1–14, <https://doi.org/10.3389/fpsyg.2019.01564> (2019).
42. McFadyen, J., Halász, V., Mermillod, M., Mattingley, J. B. & Garrido, M. I. A Rapid Subcortical Amygdala Route for Faces Irrespective of Spatial Frequency and Emotion. *J. Neurosci.* **37**, 3864–3874 (2017).
43. McFadyen, J., Mattingley, J. B. & Garrido, M. I. An afferent subcortical white matter pathway to the amygdala facilitates fear recognition. *eLife* **8**, 1–51 (2019).
44. Damasio, A. R. *et al.* Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nat. Neurosci.* **3**, 1049–1056 (2000).
45. Burra, N. *et al.* Amygdala activation for eye contact despite complete cortical blindness. *J. Neurosci.* **33**, 10483–10489 (2013).
46. Pegna, A. J., Khatib, A., Lazeyras, F. & Seghier, M. L. Discriminating emotional faces without primary visual cortices involves the right amygdala. *Nat. Neurosci.* **8**, 24–25 (2005).
47. Anderson, T. & Farb, N. A. S. Personalising practice using preferences for meditation anchor modality. *Front. Psychol.* **9**, 1–10 (2019).
48. Kirschbaum, C., Pirke, K. & Hellhammer, D. H. The “Trier Social Stress Test” - A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology.* **28**, 76–81 (1993).
49. Wang, Y., Ge, J., Zhang, H., Wang, H. & Xie, X. Altruistic behaviors relieve physical pain. *Proc. Natl Acad. Sci. U S A* **117**, 950–958 (2020).
50. Rucker, D. D., McShane, B. B. & Preacher, K. J. A researcher’s guide to regression, discretization, and median splits of continuous variables. *J. Consum. Psychol.* **25**, 666–678 (2015).
51. DeCoster, J., Gallucci, M. & Iselin, A. Best Practices for Using Median Splits, Artificial Categorization, and their Continuous Alternatives. *J. Exp. Psychopathol.* **2**, 197–209 (2011).
52. Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J. & Popovich, D. L. The median split: Robust, refined, and revived. *J. Consum. Psychol.* **25**, 690–704 (2015).
53. Tzourio-Mazoyer, N. *et al.* Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage.* **15**, 273–289, <https://doi.org/10.1006/nimg.2001.0978> (2002).

Acknowledgements

We wish to thank Paul E. Dux for helpful discussion and feedback on a manuscript draft, as well as Nicole Atcheson, Aiman Al Najjar, and Sarah Daniel at UQ’s Centre for Advanced Imaging (CAI) for help with fMRI data collection.

Author contributions

J.J.K., R.C., S.P. and J.N.K. designed research; J.J.K. performed research; J.J.K. and R.C. analyzed data; J.J.K., S.P., J.N.K. and J.R.D. wrote the paper; S.P., R.C., P.G., J.R.D. and J.N.K. provided feedback on written material.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-63846-3>.

Correspondence and requests for materials should be addressed to J.J.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020